

**UNIVERSITÀ DEGLI STUDI DI NAPOLI
"FEDERICO II"**

**DIPARTIMENTO DI SCIENZE ECONOMICHE E
STATISTICHE**



**DOTTORATO DI RICERCA IN STATISTICA
XXVIII CICLO**

**ANALISI FATTORIALE E BLOCKMODELING:
UN APPROCCIO COMBINATO PER L'ANALISI DEI
NETWORK TWO-MODE**

**FACTORIAL ANALYSIS AND BLOCKMODELING: A COMBINED
APPROACH TO ANALYZE TWO-MODE NETWORK**

Tutor:
Ch.mo Prof.re
GIUSEPPE GIORDANO

Coordinatore:
Ch.mo Prof.re
CARLO LAURO

Co-Tutor:
Ch.mo Prof.re
GIANCARLO RAGOZINI

Dottoranda:
DANIELA D'AMBROSIO

**Anno Accademico
2015/2016**

Abstract

Abstract In the present work we propose to combine Factorial Analysis and Blockmodeling techniques in order to represent and visually analyze *two-mode* networks. In social science, *two-mode* data structures are of considerable interest to the extent that they express the duality of social structure. Factorial Analysis and Blockmodeling are two of the most appropriate techniques to synthesize and graphically analyze both the underlying structure of the *two-mode* network and the relevant subsets constituting it. These techniques also preserve *two-mode* network dual nature. Our proposal aims at developing a strategy to address some of the problematic issues of Blockmodeling implementation for *two-mode* network analysis by using of Factorial Analysis tools. Due to the analytical relationships among these methodologies, we show how Factorial Methods can be a complement to *two-mode* Blockmodeling according to three main directions: as a method apt to reduce and simplify data structure; as a strategy to extract useful information in the pre-specification process; as a method to define a metric space in which to properly represent and display the positions defined through Blockmodeling, along with modes' attributes. The resulting benefits of the combination of Multiple Factorial Analysis and Blockmodeling are also shown in the analysis of heterogeneous *two-mode* network consisting of different classes of entities, with distinct sets of relationships collected for different time points, described with entities' attributes and their structural positions. This analytical approach is adopted to study a field of theatrical co-productions as part of the theatre system of Italy's Campania Region, in the framework of Bourdieu's fields theory.

Ai miei genitori

Indice

Elenco delle figure	v
Elenco delle tabelle	viii
Introduzione	x
1 Social Network Analysis e dati relazionali	1
1.1 Social Network Analysis: cenni storici e tratti distintivi	1
1.2 Forme e caratteristiche del dato relazionale	6
1.2.1 Tipi di Network	8
1.3 Rappresentazione, descrizione e analisi del dato relazionale	10
1.3.1 Elementi di Teoria dei grafi	11
1.3.1.1 Tipi di Grafo	11
1.3.1.2 Proprietà dei nodi e proprietà dei grafi	18
1.3.2 Matrici di dati relazionali: matrici di Adiacenza e di Incidenza	29
1.4 Misure di distanza per dati relazionali	33
2 Metodi di Clustering per dati relazionali: il Blockmodeling	38
2.1 Gruppi coesi e classi equivalenti	38
2.1.1 Metodi di Clustering per gruppi coesi	39
2.2 Posizioni e Sistemi di Ruoli	44
2.3 Introduzione al Blockmodeling	46
2.3.1 Equivalenza Strutturale, Equivalenza Regolare, Equivalen- za Generalizzata	47
2.4 Differenti approcci al Blockmodeling	50
2.4.1 Blockmodeling Convenzionale	52

2.4.2	Dall'approccio indiretto all'approccio diretto	57
2.4.2.1	Funzione criterio e procedura di clustering	58
2.4.3	Blockmodeling Generalizzato	60
2.5	Recenti Sviluppi	69
2.5.1	Estensione del blockmodeling a network <i>two-mode</i>	69
2.6	Blockmodeling: vantaggi e problemi aperti	73
2.7	Proposta metodologica	75
3	Blockmodeling e Analisi Fattoriale per Network <i>Two-mode</i>	80
3.1	Analisi Fattoriale e Blockmodeling: legami	80
3.2	Caratteristiche dell'Analisi delle Corrispondenze Multiple per network <i>two-mode</i>	81
3.3	Analisi delle Corrispondenze Multiple in prospettiva <i>doubling</i>	83
3.4	L'uso di punti supplementari per integrare informazioni esterne nella rappresentazione dell'ACM: Attributi e Posizioni	90
3.5	Analisi delle Corrispondenze Multiple, Equivalenza e similarità strutturale: la metrica del Chi-quadrato	94
3.6	Blockmodeling, <i>Dual-projection Approach</i> , Analisi delle Corrispondenze Multiple	96
4	Analisi delle Corrispondenze Multiple e Blockmodeling per Network <i>Two-mode</i>	102
4.1	Analisi delle Corrispondenze Multiple e Blockmodeling Esplorativo	102
4.1.1	Decomposizione di Eckart-Young e inconsistenze	104
4.1.2	Blockmodeling su matrice di affiliazione di rango inferiore	110
4.2	Analisi delle Corrispondenze Multiple e Blockmodeling Pre-specificato	124
4.2.1	Ricostruzione del contesto, selezione del numero di cluster, vincoli	125
4.3	Analisi delle Corrispondenze Multiple e Blockmodeling: combinazione	131
4.3.1	Visualizzazione e caratterizzazione delle posizioni	132
4.4	Analisi delle Corrispondenze Multiple e Blockmodeling: vantaggi	132

5	Analisi Fattoriale Multipla e Blockmodeling: un approccio combinato per l'analisi del campo	135
5.1	Analisi Fattoriale Multipla per l'analisi di Network <i>Two-Mode</i> eterogenei	135
5.2	Posizioni nella <i>Social Network Analysis</i> e nella Teoria del Campo	137
5.3	Network e campo	139
5.4	Analisi del campo attraverso l'Analisi Fattoriale Multipla e il Blockmodeling: il caso della produzione teatrale	141
5.4.1	Sistema teatrale campano: network delle co-produzioni come network <i>two-mode</i>	143
5.4.2	Network <i>Two-mode</i> delle stagioni di co-produzioni	146
5.4.3	<i>Two-mode</i> Blockmodeling delle stagioni di co-produzioni teatrali	148
5.4.4	Analisi Fattoriale Multipla e Blockmodeling: risultati . . .	167
5.5	Analisi Fattoriale Multipla e Blockmodeling: considerazioni finali	186
	Bibliografia	196

Elenco delle figure

2.1	Tipi di blocchi ideali	61
2.2	Tipi di connessioni tra blocchi	62
2.3	Blockmodeling Generalizzato	63
2.4	Blockmodeling Induttivo e Deduttivo	67
4.1	<i>Two-mode</i> Blockmodeling secondo l'equivalenza strutturale. . . .	108
4.2	Grafo bipartito con partizioni del <i>two-mode</i> blockmodel matrice artificiale. Equivalenza strutturale.	109
4.3	<i>Two-mode</i> Blockmodeling sulla matrice di affiliazione ricostruita, 1 valore singolare, \mathbf{F}_{ric_1}	112
4.4	<i>Two-mode</i> Blockmodeling sulla matrice di affiliazione ricostruita, 1 e 2 valore singolare, \mathbf{F}_{ric_2}	114
4.5	AC con <i>doubling</i> sulla matrice di affiliazione ricostruita, 1 e 2 valore singolare, \mathbf{F}_{ric_2}	116
4.6	<i>Two-mode</i> Blockmodeling sulla matrice di affiliazione ricostruita, 1, 2 e 3 valore singolare, \mathbf{F}_{ric_3}	117
4.7	AC con <i>doubling</i> sulla matrice di affiliazione ricostruita, 1, 2 e 3 valore singolare, \mathbf{F}_{ric_3}	119
4.8	<i>Two-mode</i> Blockmodeling secondo l'equivalenza strutturale. . . .	120
4.9	<i>Two-mode</i> Blockmodeling sulla matrice delle coproduzioni teatrali ricostruita attraverso i primi 4 valori singolari	123
4.10	Mappa fattoriale delle coproduzioni teatrali 2011-2012. Compagnie nello spazio delle coproduzioni.	128
4.11	Mappa fattoriale delle coproduzioni teatrali 2011-2012. Coproduzioni nello spazio delle compagnie.	129

4.12	Mappa fattoriale delle coproduzioni in quattro stagioni teatrali 2011-2012. Compagnie e posizioni delle coproduzioni nello spazio delle coproduzioni.	133
5.1	<i>Two-mode</i> network delle coproduzioni teatrali 2011-2012. <i>Degree centrality</i>	151
5.2	Modello di struttura de blockmodel co-produzioni teatrali 2011-2012.	153
5.3	<i>Two-mode</i> Blockmodeling generalizzato delle co-produzioni teatrali 2011-2012.	155
5.4	<i>Two-mode</i> network delle coproduzioni teatrali 2012-2013. <i>Degree centrality</i>	156
5.5	Modello di struttura de blockmodel co-produzioni teatrali 2012-2013.	156
5.6	<i>Two-mode</i> Blockmodeling generalizzato delle co-produzioni teatrali 2012-2013.	158
5.7	<i>Two-mode</i> network delle coproduzioni teatrali 2013-2014. <i>Degree centrality</i>	159
5.8	Modello di struttura de blockmodel co-produzioni teatrali 2013-2014.	160
5.9	<i>Two-mode</i> Blockmodeling generalizzato delle co-produzioni teatrali 2013-2014.	162
5.10	<i>Two-mode</i> network delle coproduzioni teatrali 2014-2015. <i>Degree centrality</i>	163
5.11	Modello di struttura de blockmodel co-produzioni teatrali 2014-2015.	164
5.12	<i>Two-mode</i> Blockmodeling generalizzato delle co-produzioni teatrali 2014-2015.	165
5.13	Analisi globale: rappresentazione delle compagnie nello spazio delle coproduzioni attraverso le coordinate principali.	172
5.14	Analisi globale: rappresentazione delle coproduzioni nello spazio delle compagnie per tutte le stagioni attraverso le coordinate principali.	175
5.15	Analisi globale: rappresentazione congiunta delle compagnie e delle coproduzioni.	178
5.16	Analisi globale: Analisi globale: rappresentazione delle compagnie nello spazio delle coproduzioni, attraverso le coordinate principali.	180

5.17	Analisi globale: rappresentazione delle posizioni delle compagnie e degli attributi delle compagnie e delle coproduzioni nello spazio delle coproduzioni.	182
5.18	<i>Grand table</i> \mathbb{F} per un network <i>two-mode time varying</i>	195

Elenco delle tabelle

1.1	Schema per il calcolo degli indici di Similarità	37
3.1	Matrice <i>two-mode</i> \mathbf{F}	98
3.2	Proiezione del set degli attori di \mathbf{F} tramite <i>Proportion Matching</i> .	98
3.3	Matrice <i>one-mode</i> delle co-affiliazioni <i>doubled</i> : \mathbf{FF}'	99
3.4	Matrice <i>one-mode</i> delle co-affiliazioni \mathbf{FF}'	99
4.1	Matrice <i>two-mode</i> \mathbf{F}	107
4.2	Matrice di affiliazione ricostruita, 1 valore singolare, \mathbf{F}_{ric_1}	111
4.3	Matrice di affiliazione ricostruita, 1 e 2 valore singolare, \mathbf{F}_{ric_2} . . .	113
4.4	AC con <i>doubling</i> sulla matrice di affiliazione ricostruita, 1 e 2 valore singolare, \mathbf{F}_{ric_2}	115
4.5	Matrice di affiliazione ricostruita, 1, 2 e 3 valore singolare, \mathbf{F}_{ric_3} .	116
4.6	AC con <i>doubling</i> sulla matrice di affiliazione ricostruita, 1, 2 e 3 valore singolare, \mathbf{F}_{ric_3}	118
4.7	ACM sulla matrice delle coproduzioni teatrali 2011-2012 ricostrui- ta attraverso i primi 4 valori singolari	124
4.8	ACM delle coproduzioni teatrali 2011-2012	126
5.1	Statistiche descrittive dei network e dei sub-network delle co-produzioni teatrali per le quattro stagioni.	150
5.2	AFM. Inerzia rivalutata. Analisi globale.	170
5.3	AFM. Coordinate delle stagioni sulle dimensioni dell'analisi globale.	170
5.4	AFM. Coordinate e coordinate parziali delle compagnie con con- tributi elevati.	173

5.5	Risultati dell'AFM. Coproduzioni significativamente associate alle prime quattro dimensioni.	176
5.6	AFM. Coefficiente di similarità RV per le quattro stagioni stagioni di coproduzioni.	179
5.7	Risultati dell'AFM con cluster delle compagnie e attributi. Cluster delle compagnie e attributi maggiormente associate alle prime cinque dimensioni.	184
5.8	Risultati dell'AFM con cluster delle coproduzioni e attributi. Coordinate parziali dei cluster di coproduzioni e coordinate degli attributi con contributi elevati sulle prime cinque dimensioni.	185

Introduzione

Uno degli sforzi conoscitivi più rilevanti nella SNA è legato allo sviluppo di metodi che consentano l'analisi e la comprensione della struttura dei modelli di connessioni alla base di un network sociale e delle sue conseguenze. L'analisi dei modelli di connessioni sono fondamentali per la comprensione delle azioni individuali e del funzionamento del sistema sociale nel suo complesso. All'interno di queste elaborazioni assumono particolare rilievo i metodi che consentono l'analisi delle strutture di dati *two-mode* (o multi-mode) che racchiudono la dualità della struttura sociale⁽⁴³⁾, ossia la dinamica individuo/collettività espressa attraverso i legami di appartenenza, o affiliazione, che connettono gli attori a differenti livelli della struttura sociale (individuo/ gruppo/organizzazione/ comunità). Le tecniche di Analisi Fattoriali, in particolare l'Analisi delle Corrispondenze, e le tecniche di Blockmodeling sono tra gli strumenti più appropriati per sintetizzare, analizzare e rappresentare graficamente sia la struttura sottostante un network *two-mode* sia i sottoinsiemi significativi che la compongono, preservandone la natura duale. La seguente proposta si presenta come sviluppo di una strategia che consenta di affrontare alcune delle problematiche legate all'applicazione delle tecniche di Blockmodeling per l'analisi dei network *two-mode*. Tale strategia prevede il ricorso all'ausilio degli strumenti offerti dalle tecniche di Analisi Fattoriale. L'approccio sviluppato è vicino, ma non completamente incorporato, in quello che è stato definito come approccio posizionale. In esso, l'obiettivo principale è quello di svelare la struttura sottostante il network identificando sottoinsiemi di unità e di relazioni rilevanti nella struttura relazionale, quali posizioni e ruoli. In questo contesto, l'approccio adottato cerca di combinare i vantaggi principali sia delle tecniche di Analisi Fattoriale sia di quelle Posizionale che, pur provenendo da tradizioni matematiche e statistiche differenti, presentano inte-

ressanti punti di connessione. Entrambe cercano di individuare e sintetizzare la struttura sottostante i fenomeni osservati guardandoli da una prospettiva globale. Entrambi i metodi sono considerati compatibili e complementari in quanto considerano lo stesso oggetto con diversi livelli di dettaglio e da diversi punti di vista, consentendone un'analisi approfondita. Seguendo un approccio diretto, si sceglie di analizzare e rappresentare i network *two-mode* attraverso l'Analisi delle Corrispondenze Multiple (ACM) con codifica disgiuntiva completa adottando la prospettiva del Doubling (70, 71, 119). La scelta di questa versione dell'ACM è legata alle sue specifiche caratteristiche formali che la rendono particolarmente appropriata al trattamento dei network *two-mode*.

Nell'ambito dell'applicazione del Blockmodeling ai network *two-mode*, che costituisce la struttura di riferimento negli sviluppi della presente proposta metodologica, è possibile individuare alcune problematiche: (i) Dimensioni elevate del network che comportano soluzioni più instabili con più partizioni plausibili (*well fitting*), elevati valori della funzione criterio e difficile identificazione sia del numero delle posizioni sia dei componenti al loro interno, con un incremento del carico computazionale combinatorio(77, 81, 83). (ii) Interpretazione e spiegazione dei fenomeni di network. Utilizzo privilegiato di dati relazionali e proprietà strutturali (misure relative alle caratteristiche del network) a discapito degli attributi descrittivi delle entità (113). (iii) Difficoltà di operare una scelta tra partizioni ugualmente ben adattate senza ulteriori informazioni, sia teoriche sia empiriche, sul fenomeno sociale in esame. (iv) Necessità di informazioni sostantive o empiriche che orientino le scelte legate alla pre-specificazione(81, 83). (v) Visualizzazione dei dati relazionali. I risultati del *two-mode* Blockmodeling possono essere rappresentati mediante grafo ridotto (grafo immagine), ma senza la possibilità di valutare la distanza tra le posizioni.

Nell'analisi dei *two-mode* network, l'ACM può essere di supporto al *two-mode* Blockmodeling secondo tre principali direzioni: come metodo di riduzione e semplificazione della struttura dei dati; come strategia per estrarre informazioni utili nel processo di pre-specificazione; come metodo per definire uno spazio metrico appropriato in cui rappresentare e visualizzare le posizioni definite attraverso il Blockmodeling. Dal canto suo, il Blockmodeling fornisce una partizione simultanea, ma distinta, delle righe e delle colonne della matrice *two-mode*, consentendo

di mettere in relazione diretta le unità di un set con quelle dell'altro set dalle quali sono definite. Inoltre, le rappresentazione ridotte ottenute tramite il Block-modeling, ossia il grafo ridotto e la matrice immagine, consentono di visualizzare in modo sintetico le posizioni e le relazioni tra le posizioni (blocchi). Pertanto, l'integrazione di entrambe le tecniche offre l'opportunità di valorizzare i singoli contributi informativi e migliorare l'interpretazione globale dei risultati con riflessi positivi sulla capacità di comprensione del fenomeno sociale in esame. Con tali premesse, la presente proposta può essere considerata come un approccio multi-step in cui le due tecniche possono interagire in una o più fasi del processo di analisi dei dati, a seconda delle specifiche esigenze.

Le opportunità derivanti dalla combinazione di tecniche di Analisi Fattoriale e tecniche di Blockmodeling sono visibili anche in una prospettiva più ampia legata alla gestione di network eterogenei. I fenomeni sociali sono complessi, multidimensionali e pertanto la loro comprensione richiede la capacità di mettere in relazione aspetti diversi a differenti livelli di analisi. Ne consegue la necessità di trattare informazioni molteplici e di natura diversa, spesso, in un'ottica diacronica. Nel tentativo di compiere un passo in questa direzione si propone una strategia di analisi che combini l'Analisi Fattoriale Multipla (AFM) (87, 169) e il Blockmodeling per lo studio dei network *two-mode* eterogenei, in cui sono presenti dati relativi ai legami tra due classi di entità differenti, rilevati in momenti temporali diversi, che includono gli attributi descritti delle entità e i modelli di relazioni che intercorrono tra loro (le posizioni strutturali). Tale approccio analitico sarà adottato per lo studio del campo della co-produzione teatrale nell'ambito del sistema teatrale regionale campano (182, 183), nella cornice teorica della teoria del campo di Bourdieu (35).

Al fine di illustrare il percorso concettuale e metodologico in cui si articola la proposta, la tesi è strutturata nelle seguenti tappe. Il primo capitolo illustra i tratti costitutivi e i concetti essenziali della prospettiva della Social Network Analysis. In esso, vengono ricostruite le radici teoriche dell'approccio; le forme assunte dalle strutture di dati relazionali e le loro caratteristiche centrali. Ampio spazio viene dedicato alla definizione degli strumenti utilizzati per la rappresentazione, descrizione e analisi del dato relazionale, che saranno impiegati nel corso

della trattazione. Infine, una particolare attenzione è dedicata alle misure di distanza e alle caratteristiche assunte da questo concetto nel contesto della teoria dei grafi, con relative implicazioni in termini di valutazione dello spazio tra i nodi nelle rappresentazioni grafiche dei dati relazionali. Il secondo capitolo è dedicato ai metodi di clustering per dati relazionali. In esso vengono ricostruiti i principali approcci e metodi utilizzati per l'individuazione di sotto-strutture rilevanti nel network, sia in termini di gruppi coesi sia di classi equivalenti. Queste ultime costituiscono un punto di interesse focale del presente lavoro. In particolare, vengono mostrate le radici concettuali e le caratteristiche dell'approccio posizionale e, nello specifico, del metodo del Blockmodeling, che comprende una famiglia di tecniche dedicate alla partizione del network sulla base di un certo tipo di equivalenza. Vengono presentati i vari approcci al Blockmodeling e, di questi, sono esaminati dettagliatamente il Blockmodeling generalizzato e la sua estensione per l'analisi dei network *two-mode*. Infine, ne vengono mostrati i principali vantaggi e le criticità che costituiscono la base per la formulazione della proposta metodologica con cui il capitolo si conclude e che viene illustrare nei capitoli successivi. Il terzo capitolo è dedicato alla ricostruzione dei legami concettuali e analitici tra il Blockmodeling e l'Analisi Fattoriale sui quali si basa la scelta di combinare i due approcci per l'analisi dei network *two-mode*. In particolare, vengono presentate le caratteristiche dell'Analisi delle Corrispondenze Multiple nell'analisi dei network *two-mode*; del particolare contributo offerto dall'adozione della prospettiva del doubling; delle caratteristiche tecniche che permettono l'integrazione dei risultati del Blockmodeling in un spazio metrico. Inoltre, vengono illustrati i dettagli delle relazioni tra equivalenza e similarità strutturale sulla base delle proprietà assunte dalla metrica del Chi-quadrato nell'Analisi delle Corrispondenze Multiple. Infine, vengono mostrate le connessioni dei metodi proposti per la definizione della similarità strutturale con altri presenti in letteratura nell'ambito dell'analisi di Blockmodeling. Nel quarto capitolo viene presentato nel dettaglio l'utilizzo della dell'Analisi Fattoriale come strumento di supporto all'analisi di Blockmodeling che costituisce la proposta metodologica del presente lavoro. Nello specifico, vengono riportati i risultati relativi alla strategia di riduzione e sintesi della struttura del network attraverso il metodo di decomposizione di Eckart-Young (85) con i relativi effetti in termini di Blockmodeling; i risultati derivanti dall'utilizzo dell'Analisi

delle Corrispondenze Multiple come metodi di estrazione di informazioni utili nel processo di pre-specificazione nell'ambito del Blockmodeling pre-specificato; in ultimo, i risultati ottenuti dall'impiego dell'Analisi delle Corrispondenze Multiple per la visualizzazione e caratterizzazione delle posizioni. Nel quinto e ultimo capitolo vengono mostrati i vantaggi conoscitivi derivanti dalla combinazione dei metodi di Analisi Fattoriale e Blockmodeling per l'analisi dei network eterogenei. Nello specifico vengono presentati i risultati relativi all'impiego dell'approccio dell'Analisi Fattoriale Multipla come metodo per gestire network *two-mode* eterogenei che variano nel tempo e che presentano informazioni aggiuntive, quali attributi descrittivi e posizioni strutturali derivanti dall'analisi di Blockmodeling. Le potenzialità derivanti dell'adozione di tale approccio vengono mostrate attraverso un'applicazione dell'Analisi Fattoriale Multipla e Blockmodeling per l'analisi del campo delle co-produzioni teatrali parte del sistema teatrale della Regione Campania, in Italia, nel quadro concettuale della teoria dei campi di Bourdieu. Nel corso del capitolo viene presentato il quadro teorico di riferimento e i principali risultati dell'analisi.

Capitolo 1

Social Network Analysis e dati relazionali

1.1 Social Network Analysis: cenni storici e tratti distintivi

Gli sviluppi teorici e metodologici della Social Network Analysis (SNA) hanno profonde radici negli studi della tradizione sociologica e psicosociale realizzati a partire dagli anni '30. In questo periodo i ricercatori iniziano ad utilizzare il concetto di "network sociale" per rappresentare e descrivere la struttura sociale (176, 177, 178). Tra i riferimenti centrali si distingue la "sociologia formale" del sociologo tedesco Simmel, interessata alle proprietà formali delle relazioni interpersonali. Tali proprietà vengono considerate alla base dello studio delle configurazioni delle relazioni sociali, generate dall'intreccio delle interazioni reciproche. Accanto alla sociologia tedesca, si riconoscono gli influssi della teoria della "Gestalt", sviluppatasi nell'ambito della psicologia americana, visibile soprattutto nei contributi di Lewin (140) e Moreno (155). Quest'ultimi sono tra i primi ad introdurre concetti quali "campo" o "spazio" delle relazioni sociali e a porre l'attenzione sulle caratteristiche delle relazioni (177). Le riflessioni di questi teorici hanno un notevole impatto sugli sviluppi teorico-concettuali della SNA poichè, da un lato, sottolineano l'influenza delle relazioni di gruppo sugli individui e, dall'altro, impiegano gli strumenti e la formalizzazione matematica per descrivere le relazioni. In particolare, Moreno (155) viene considerato il primo ad utilizzare il sociogramma come strumento per la rappresentazione delle forme delle relazioni interpersonali, o meglio, come una delle tecniche per analizzare l'evoluzione, l'organizzazione di gruppi e la posizione degli individui al loro interno. Tale compito

costituisce l'oggetto di studio di quella disciplina che da lui prese il nome di "sociometria" (102). Dal canto suo, Lewin (140), nell'ambito della teoria del campo, costituisce il primo ad utilizzare strumenti matematici per descrivere la relazione tra gruppo e ambiente, gettando le basi per gli sviluppi successivi di Cartwright e Harary (60) in cui i comportamenti di gruppo sono stati analizzati con l'ausilio della teoria dei grafi (102, 165). Un notevole influsso, inoltre, proviene dagli studi classici di Warner e Lunt (190), sulla città di "Yankee City", e Mayo (146), sulla compagnia elettrica Hawthorne di Chicago. Tali studi, attraverso considerazioni empiricamente fondate, misero in luce l'importanza delle relazioni informali interpersonali nella creazione di sottogruppi coesivi (*clique*) in grado di contribuire all'integrazione degli individui nella struttura organizzativa e comunitaria. Questi studi, insieme a quelli della tradizione psicosociale, furono alla base di ulteriori sviluppi tecnici in direzione dei metodi matriciali per la rappresentazione e analisi della struttura relazionale dei sistemi sociali complessi (165, 177).

A partire da queste tradizioni di ricerca, negli anni '50 i principali sviluppi nell'ambito della SNA si hanno ad opera di un gruppo di antropologi riuniti intorno a Gluckman noti come "Scuola di Manchester". Il loro lavoro si fonda su una visione critica dell'impostazione struttural-funzionalista classica riconducibile a Radcliffe-Brown fondato sulla ricerca dei meccanismi di equilibrio strutturale e sul concetto di gruppo corporato come elemento costitutivo della società. Quello degli antropologi di Manchester è un approccio analitico e processuale, volto a cogliere i meccanismi di conflitto e di cambiamento. Il concetto di network sociale è adottato come superamento di quello di gruppo corporato, inadeguato a descrivere la forma mutevole delle relazioni sociali. L'unità di analisi è l'individuo, con i suoi mutevoli legami, che viene concepito come capace di manipolare e, allo stesso tempo, essere manipolato dai suoi legami (165). Barnes (5) è stato il primo ad utilizzare il concetto di network di relazioni, ulteriormente approfondito grazie alle ricerche sul campo della Bott(33) sulle reti di parentela. Tra questi ricercatori, Mitchell (153) fu il primo a compiere un'opera di sintesi e di sistematizzazione della metodologia della SNA.

Accanto a questa impostazione di ricerca, a partire dagli anni '70, viene ad affermarsi quella dei sociologi americani, raccolti intorno a White, nota con il nome di "Scuola di Harvard". Questi studiosi iniziarono a orientare lo studio dei network

sociali in direzione di un'analisi strutturale fino a sviluppare una metodologia formale per SNA. Alla base delle riflessioni di questo approccio vi è il recupero del concetto di struttura come elemento fondamentale a partire dal quale è possibile spiegare il comportamento degli individui. L'unità di analisi, quindi, si sposta sulla relazione e i fenomeni sociali possono essere compresi a partire dall'analisi dei modelli di relazioni significative esistenti tra gli elementi del sistema (165). Numerosissimi sono i contributi che seguono questa impostazione e che, nel corso del tempo, hanno concorso a rafforzare la connessione tra sviluppi teorico-concettuali e loro formalizzazioni matematiche (22, 102). In questa visione, infatti, l'operazionalizzazione di ogni concetto in termini relazionali è radicato in una nozione teoricamente rilevante per le sue implicazioni sociali; ne sono esempi la misura di densità del network, che esprime matematicamente il concetto di coesione, o la formalizzazione dell'equivalenza strutturale, espressione matematica del concetto di ruolo (143). Il sorgere di una nuova visione del sistema sociale come "network", in cui la sua struttura viene concepita come "modello regolare e persistente nel comportamento delle parti elementari di un sistema sociale" (22, p. 1), ha portato i ricercatori a sviluppare e utilizzare strumenti matematici, come la teoria dei grafi, per concettualizzare, descrivere e analizzare tali modelli di connessione e le loro implicazioni.

L'incremento degli sviluppi in questa direzione sono stati consentiti da due innovazioni matematiche: lo sviluppo di interpretazioni algebriche dei network sociali e l'utilizzo delle tecniche di Scaling Multidimensionale (MDS). Centrali sono i contributi di White che applica i concetti algebrici per rappresentare le strutture di parentela (196); Laumann (135) che utilizza i metodi di Scaling Multidimensionale per tradurre graficamente le differenze sociali in "distanze sociali"; di Boorman e White (23) che sviluppano i metodi di analisi matriciali per lo studio delle posizioni sociali. A partire dai contributi di questo gruppo di ricerca l'approccio della SNA si è diffuso negli altri continenti fino ad essere adottato anche in altre discipline al di fuori delle scienze sociali, come in fisica, dove il concetto di network si è rivelato particolarmente utile. Infatti, a partire dagli anni '90, si è avuta una esplosione di contributi da parte dei fisici che hanno iniziato ad applicare il concetto di network ai fenomeni sociali. L'emergere di questa area di

studi, nota come "fisica sociale", si fonda sui lavori di Watts e Strogatz (193) che riprendono alcune idee di Milgram (152) sul concetto di "*small world*" (177, 178).

Il concetto di network si è diffuso in tutte le discipline *network-based* (132), il cui oggetto di studio può essere concepito come "sistema di parti interconnesse" (strutture sociali, strutture molecolari, sistemi di comunicazione, ecc.). Di conseguenza, la SNA è divenuta parte di un programma di ricerca più ampio, noto come Network Analysis, caratterizzato da contributi provenienti da discipline diverse (scienze sociali, scienze matematiche e statistiche, scienze fisiche, biologia, ecc.) che hanno come elemento aggregante lo sviluppo e la condivisione di un insieme di concetti e metodi per la misura, rappresentazione e analisi dei fenomeni.

Fatte queste elementari premesse è possibile notare come l'approccio della SNA si presenti con alcuni caratteri innovativi rispetto a quello individualista o "atomistico" (69) tradizionale che fonda le sue spiegazioni dei fenomeni sociali a partire dagli attributi descrittivi dei singoli soggetti. Rispetto a quest'ultimo, si possono individuare alcuni tratti distintivi della prospettiva della SNA (192, 194):

- gli attori e le loro azioni sono considerati come interdipendenti e non come unità autonome;
- i fenomeni sociali sono concepiti come processi relazionali o risultati strutturali;
- i legami sono visti come veicolo di flussi di risorse sia materiali sia immateriali;
- le condizioni strutturali del network costituiscono sia opportunità sia vincoli per l'azione individuale;
- i modelli di network concettualizzano la struttura come modelli stabili di relazioni tra attori.

In questa ottica, quindi, la SNA, e nello specifico l'approccio strutturale, si pone l'obiettivo di delineare i modelli stabili di relazioni tra le entità interagenti e le implicazioni di questi legami sui fenomeni sociali (192, 194). L'oggetto di studio sono i network sociali definiti come un set di entità, individuali o collettive,

connesse tra loro da uno o più tipi di relazioni. L'assunto alla base dei suoi strumenti concettuali, metodologici e analitici è che il comportamento sociale possa essere concettualizzato e analizzato in termini di strutture di relazioni tra gli attori del sistema piuttosto che in termini di attributi degli attori (176, 178, 192).

Ripercorrendo lo sviluppo della disciplina, si possono individuare quattro elementi caratterizzanti la SNA (102, 106):

- (i) radicamento sull'intuizione strutturale che attribuisce un ruolo centrale ai legami sociali tra gli attori;
- (ii) raccolta e analisi di dati empirici che codificano le relazioni sociali che legano gli attori;
- (iii) ricorso alle visualizzazioni grafiche per rappresentare i modelli di legami;
- (iv) sviluppo di modelli matematici e computazionali per descrivere e spiegare questi modelli.

A queste, possono essere aggiunte anche (18):

- (v) centralità di una comprensione dei fenomeni basata su considerazioni sostanziali e concettuali profonde;
- (vi) rilevanza della dimensione dinamica dei processi sociali di creazione, mantenimento e dissolvimento del network sociale;
- (vii) rilevanza dello studio del contesto in cui il network sociale è inserito.

In ultima istanza, è importante sottolineare come, nel corso del tempo, siano state mosse numerose critiche all'approccio della SNA. Prima di tutto, essa viene considerata come eccessivamente focalizzata sulle descrizioni dei fenomeni a discapito della teorizzazione. Inoltre, le viene contestato l'eccessivo ricorso alla formalizzazione matematica e un accento quasi esclusivo sulla struttura piuttosto che sui contenuti veicolati e sulla capacità di *agency* degli attori. Infine, le si attribuisce un approccio statico ai fenomeni, poco attento alle dinamiche del cambiamento e la produzione di spiegazioni che sono radicate nel network ma

mancono del contesto in cui tale network è calato. Alcuni studiosi, a tal proposito, fanno uno sforzo di chiarificazione cercando di mettere in luce come alcuni di questi punti critici abbiano riguardato soprattutto gli albori della SNA e come, nel corso del tempo, i nuovi sviluppi si siano mossi, e si stiano muovendo, in direzione di un loro approfondimento (32).

1.2 Forme e caratteristiche del dato relazionale

Come si è detto, l'obiettivo delle analisi della SNA è quello di individuare i modelli stabili di relazioni tra le entità interagenti e le implicazioni sul comportamento e sulle scelte delle entità, a differenti livelli di analisi (intera struttura sociale; attore individuale; diade, triade, subset) (144, 192, 194). In tale approccio l'unità di analisi non è l'individuo in sé ma la diade, o meglio il legame tra coppie di entità, siano esse singoli individui o aggregati di individui. Pertanto, le variabili rilevate sono principalmente di carattere strutturale, ossia le relazioni. Accanto ad esse vengono considerati anche gli attributi delle entità, tuttavia, assumono un ruolo secondario.

Uno dei punti critici della SNA è stata la mancanza di una chiara definizione di una tipologia di legami (32). Questa mancanza, talvolta, ha portato alla confusione tra tipi di legami, come stati di relazione, e le loro funzioni. Nel tentativo di superare tale ambiguità in letteratura viene proposta la distinzione tra: (i) relazioni di somiglianza; (ii) relazioni sociali; (iii) interazioni; (iv) relazioni di flusso (32).

Le relazioni di somiglianza sono caratterizzate dalla condivisione di un attributo (stesso orientamento politico, caratteristiche demografiche, atteggiamenti), dalla co-appartenenza ad un gruppo o dalla co-partecipazione ad attività. Le relazioni sociali sono costituite dalle relazioni di parentela (padre, fratello), di ruolo (amico, datore di lavoro, studente), legami affettivi (simpatie o antipatie), relazioni di percezione (conoscere) ¹. Le interazioni sono basate sul comportamento

¹Le relazioni sociali hanno in comune con quelle di similarità la durata, che si protrae fino ad esistenza del legame (31, 32)

legato a specifiche azioni (conversazione, aiuto) ¹. Le relazioni di flusso sono basate sugli scambi, o transazioni (di risorse, informazioni). Questi possono includere le relazioni tramite le quali risorse, informazioni o influenza attraversano il network (31, 32, 143).

Scelto il tipo, o i tipi, di relazione sul quale rilevare le informazioni, tali dati possono essere raccolti per tutti i legami che connettono le unità della popolazione di interesse (*Whole-network*) oppure per il set di legami connessi ad un attore focale (*Ego-network*). I dati relativi ai legami di network possono essere rilevati tramite *survey*, questionari, interviste, archivi, osservazione, diari, esperimenti (144, 145). Al termine della raccolta, i dati relazionali possono essere rappresentati, descritti e analizzati attraverso alcuni strumenti matematici che costituiscono una parte fondamentale della prospettiva della SNA. Quest'ultima, consentendo molteplici livelli di analisi (network totale/ sub-strutture/ network ego-centrati) rende evidente come, da un lato, le differenze tra attori siano definite dai vincoli e dalle opportunità derivanti dal modo in cui essi sono incorporati nel network. Dall'altro, sottolinea come la struttura del network e le sue trasformazioni siano radicate e veicolate dalle interazioni locali tra gli attori (124).

Come osservato in precedenza, all'interno della SNA vi è uno stretto legame tra concetti teorici relativi ai fenomeni sociali e le misure e i metodi statistici sviluppati per rilevarli e analizzarli. I principali indici statistici utilizzati per descrivere le proprietà strutturali del network ruotano intorno ad alcune idee centrali: cogliere, ai vari livelli, il grado e i meccanismi di integrazione e differenziazione strutturale. Questi obiettivi, sono perseguiti sulla base di concezioni diverse dei modelli di struttura del network che, sostanzialmente, riflettono approcci diversi all'analisi dei fenomeni di network. In particolare, questi approcci, pur condividendo l'obiettivo di descrivere il network delle relazioni, differiscono nel quadro di riferimento all'interno del quale un attore viene analizzato e nei meccanismi esplicativi dei fenomeni in cui è coinvolto (29, 56).

¹Ciò che distingue le interazioni dalle relazioni sociali è il loro carattere frammentario, discontinuo. Una relazione sociale dura fino a quando continua ad esistere la condizione che l'ha generata, mentre l'interazione dura il tempo dell'azione specifica, come collaborare ad un progetto o andare a teatro insieme (31, 32). Inoltre, le interazioni spesso si verificano nel contesto delle relazioni sociali e talvolta, negli studi di network, sono utilizzate come *proxy* l'una dell'altra (143).

Principalmente, si possono distinguere un approccio relazionale (o connessionista) e un approccio posizionale (o strutturalista). Nell'approccio relazionale i modelli di network descrivono l'intensità della relazione tra coppie di attori. Nell'approccio posizionale, invece, essi descrivono il modello delle relazioni che definisce la posizione di un attore in un sistema di attori. L'approccio relazionale descrive la struttura del network in termini di relazioni in cui sono coinvolti gli individui e nella misura in cui gli attori sono connessi all'interno di gruppi primari coesi come le *clique*. L'approccio posizionale, invece, descrive tale struttura come interconnessa, differenziata e stratificata in base all'importanza, allo status e al set di ruoli degli attori che la compongono (29, 56).

1.2.1 Tipi di Network

A partire dalla natura delle classi di attori e dalle proprietà dei legami esistenti tra loro è possibile individuare diversi tipi di network sociali (192). Questa distinzione risulta indispensabile ai fini di una corretta scelta delle tecniche di analisi. Si indica con il termine "mode" del network il numero delle classi di entità su cui vengono misurate le variabili strutturali. Nei network *one-mode*, il tipo più diffuso, è presente un'unica classe di attori; nei network *two-mode* sono presenti due classi differenti di unità, due classi di attori diversi o una di attori e una di eventi; nei network *tree-mode* o *multi-mode* sono presenti tre o più classi di unità differenti. A queste vanno aggiunte strutture di dati come gli *egocentric-network*, in cui sono presenti i legami che uno specifico soggetto, *Ego*, ha con altri soggetti, *alters*, e i legami che questi ultimi hanno tra loro. Infine, è possibile individuare altri tipi di network speciali(192).

Più nel dettaglio, un network *one-mode* è caratterizzato da un unico set di unità o attori su cui vengono rilevate le informazioni. Gli attori possono essere di natura diversa: persone, gruppi, organizzazioni, nazioni. Le relazioni sono misurate sulla coppia di attori e riguardano specifici tipi di connessioni, specifici tipi di legami, che rispondono agli interessi di ricerca. Oltre alle relazioni, che costituiscono "*the building blocks of network analysis*" (131, p. 12), i dati sul set di attori possono contenere anche misure sulle caratteristiche degli attori, del tutto analoghe alle variabili analizzate negli approcci tradizionali.

Un network *two-mode* è costituito da due distinte classi di unità, che possono essere due classi di attori diversi o una di attori e una di eventi.

Nel primo caso possono essere definiti come network *two-mode* diadici (*dyadic two-mode network*). In essi la relazione è misurata su una coppia di attori che appartengono alle due classi. Anche in questi network i tipi di attori possono essere di diversa natura e le informazioni rilevate possono riguardare sia il tipo di relazione sia, eventualmente, gli attributi relativi ad ognuna delle classi di attori. Un aspetto specifico dei network *two-mode* diadici è che almeno una relazione deve essere misurata tra i due differenti set. Possono esserci casi di network *two-mode* in cui sono rilevate anche relazioni tra attori della stessa classe, ma l'aspetto vincolante è che almeno una relazione intercorra tra i due set.

Il secondo tipo di network *two-mode* è costituito da una classe di attori e una di eventi e prende il nome di network di affiliazione (*affiliation* o *membership network*). In questo tipo di network sociale la relazione, ossia il legame, è definito dalla partecipazione, o affiliazione, ad un'attività o ad un evento appartenente all'altra classe di entità. In questo caso, quindi, il primo modo si riferisce alla classe di attori e il secondo modo a quello degli eventi. Questo tipo di network non è diadico in quanto la relazione è rilevata su subset di attori. Ciò che caratterizza questo tipo di network è la condizione che ogni attore partecipi o sia affiliato a una, o più, attività o eventi. Nei network di affiliazione gli attori della prima classe sono in "relazione" attraverso la partecipazione congiunta agli eventi. Gli eventi possono essere definiti in base all'appartenenza ad un'organizzazione o la partecipazione ad un evento sociale o ad un'attività. La natura dell'evento dipende dall'attore coinvolto che può essere sia singolo (membro di comitati) sia collettivo (nazioni che appartengono a trattati). Anche in questo caso, oltre alle relazioni, è possibile rilevare le caratteristiche descrittive di entrambe le classi, attori ed eventi(192).

Una struttura di dati più complessa è costituita dai network multilivello, che possono essere definiti come network composti da network *one-mode* per ciascun livello e un network *two-mode* che unisce le classi di entità di livelli differenti. Un network multilivello può essere anche multi-relazionale ma è necessario tenere distinti i due concetti. Le entità in un network multilivello sono composte da

diversi tipi di unità, ed ogni tipo corrisponde ad un livello. Un network multirelazionale è un network in cui diverse relazioni vengono misurate su uno (o più) classi di entità. In un network multilivello, sono presenti relazioni sia tra le unità di ogni singolo livello (nei due *one-mode*), sia tra le unità di livelli differenti (network *two-mode*). Ad esempio, un network multilivello composto da due livelli, è una struttura in cui sono presenti N unità di livello inferiore e M unità di livello superiore, nello specifico: a) un network *one-mode* di individui o primo livello; b) un network *one-mode* di istituzioni o secondo livello; c) un network *two-mode* di affiliazione che lega gli individui alle istituzioni, quindi unisce le unità del primo e secondo livello (206, 207) .

1.3 Rappresentazione, descrizione e analisi del dato relazionale

I dati relazionali possono essere rappresentati, descritti e analizzati attraverso gli strumenti matematici della teoria dei grafi e dell'algebra matriciale.

La teoria dei grafi (126) rappresenta il cuore della SNA, è costituita da una serie di teoremi che consentono la formalizzazione, rappresentazione e analisi delle proprietà formali di un sistema composto di punti e linee. Grazie agli assunti di questa disciplina le strutture relazionali possono essere concettualizzate e rappresentate come sistemi di punti e linee, ossia i nodi, per rappresentare le unità, e i legami, per le connessioni tra le unità. La teoria dei grafi fornisce le basi per l'analisi delle proprietà formali dei network sociali (o sociogrammi).

I dati relazionali possono essere rappresentati anche in forma matriciale (126) (spesso chiamata matrice sociometrica) che costituisce il primo passo per il calcolo degli indici di network. Il grafo del network sociale e la matrice possono contenere le stesse informazioni, le righe e le colonne corrispondono ai punti e i valori nelle celle alla presenza, o assenza, del legame tra le coppie di punti. L'organizzazione in matrice offre il vantaggio di poter analizzare direttamente le proprietà formali, ricorrendo alle operazioni dell'algebra matriciale, il che rappresenta un notevole vantaggio in presenza di reti di grandi dimensioni (178, 192).

I principi della teoria dei grafi possono essere estesi per tenere conto delle caratteristiche dei legami (direzione, valenza, intensità, ecc.) dando origine ad un

"Network". Il concetto di network in matematica viene utilizzato per indicare questo tipo di grafo di conseguenza *"la social network analysis è l'analisi dei sistemi di relazioni sociali rappresentate attraverso i network"*(178, p. 4).

1.3.1 Elementi di Teoria dei grafi

Un network sociale può essere descritto in modo formale ricorrendo alle notazioni convenzionali offerte dalla teoria dei grafi. Un grafo è costituito da un insieme di punti uniti tra loro da linee (199). La rappresentazione di un grafo non tiene conto delle regole spaziali di tipo euclideo e, quindi, non considera le proprietà metriche della rappresentazione (62, 192). Il ricorso agli assunti della teoria dei grafi consente di ottenere numerosi vantaggi.

Il linguaggio matematico offre strumenti che consentono la concettualizzazione, definizione, misurazione e analisi di molte proprietà strutturali sociali. Inoltre, esso consente di sviluppare, verificare e visualizzare modelli teorici sulle strutture sociali (192).

La teoria dei grafi fornisce gli strumenti concettuali di base della SNA ¹. Essa, infatti, offre una rappresentazione di un network sociale come modello di un sistema sociale costituito da un insieme di attori e dai legami tra loro. Ovviamente, proprio in quanto modello, si tratta di una rappresentazione semplificata, che sintetizza una parte degli elementi di una situazione concreta.

1.3.1.1 Tipi di Grafo

Grafi semplici o non diretti. Un grafo semplice o non diretto (*simple graph* o *undirected graph*) è un modello di network sociale con relazioni non-dirette binarie. In esso vi è solo l'informazione relativa alla presenza o assenza del legame. Relazioni non direzionate o bidirezionali includono le co-appartenenze, i legami di parentela, le interazioni ecc.

In termini formali, un grafo \mathcal{G} è costituito da due insiemi di oggetti, un insieme di nodi, chiamati anche punti o vertici, che rappresentano gli individui, e un insieme di linee, note anche come legami o archi, che rappresentano le connessioni tra

¹In questo contesto ci si riferirà alla notazione utilizzata da Wasserman e Faust (192) e Wilson (199) .

coppie di punti. Si definisce \mathcal{V} il set che contiene n nodi, $\mathcal{V} = \{v_1, v_2, \dots, v_n\}$ e \mathcal{E} è il set che contiene l linee tra coppie di nodi, $\mathcal{E} = \{e_1, e_2, \dots, e_l\}$. In un grafo semplice ogni linea è una coppia non ordinata di nodi, $e_k = (v_i, v_j)$. Dato che le linee sono coppie di nodi non ordinate, la linea e_k tra il nodo v_i ed v_j è identica alla linea tra v_j e v_i , $e_k = (v_i, v_j) = (v_j, v_i)$.

In queste coppie di nodi non ordinate, ogni linea può collegare una coppia di nodi una sola volta e sono escluse le linee che collegano un nodo con se stesso (v_i, v_i) (*loops* o *reflexive ties*). Quindi, si definisce grafo semplice (*simple graph*) il grafo in cui non ci sono *loops*, le linee sono collegamenti senza direzione tra coppie di nodi, e per ciascuna coppia di nodi non può passare più di un legame.

Formalmente si definisce come grafo \mathcal{G} l'insieme del set di nodi \mathcal{V} e il set di linee \mathcal{E} , $\mathcal{G}(\mathcal{V}, \mathcal{E})$. Due nodi v_i ed v_j si dicono adiacenti se la linea che li unisce $e_k = (v_i, v_j)$ fa parte del set delle linee \mathcal{E} . Un nodo si dice incidente con una linea, e una linea incidente con un nodo, se è parte della coppia di nodi non ordinata che definisce la linea. A partire da questi è possibile definire come banale (*trivial*) un grafo composto da un unico nodo e vuoto (*empty*) un grafo composto da nodi non legati tra loro, quindi senza linee.

In un grafo semplice, $\mathcal{G}(\mathcal{V}, \mathcal{E})$, il suo ordine (*order*) è dato dal numero dei suoi punti e la sua dimensione (*size*) è data dal numero delle sue linee. Si noti che la posizione dei punti sul piano è arbitraria, pertanto non rispetta le regole dello spazio metrico euclideo rendendo priva di significato la lunghezza delle linee tra i punti.

Grafi diretti. Nei grafi semplici le relazioni non hanno direzione, nelle situazioni concrete, invece, molte relazioni sono dotate di direzione (*directional*) dando origine ad un altro tipo di grafo molto utilizzato nelle applicazioni di SNA. Sono esempi di relazioni dirette quelle in cui il legame è orientato da un attore ad un altro (import/export di beni, scelte di amicizia, ecc.). In una relazione diretta l'attore i può scegliere l'attore j , sebbene questo non implichi necessariamente che l'attore j scelga l'attore i .

Una relazione diretta può essere rappresentata attraverso un grafo diretto (*digraph*). Un grafo diretto è costituito da un insieme di nodi, che rappresentano gli attori in un network, e un insieme di archi diretti tra coppie di nodi, che

rappresentano legami diretti tra attori. Un grafo diretto si differenzia da quello semplice per il fatto che, in esso, viene specificata la direzione delle linee. Queste linee orientate sono chiamate archi (*arc*).

Un grafo diretto $\mathcal{G}_d(\mathcal{V}, \mathcal{E})$ è formato da un insieme di nodi $\mathcal{V} = \{v_1, v_2, \dots, v_n\}$ e da un insieme di archi $\mathcal{E} = \{e_1, e_2, \dots, e_l\}$. Ciascun arco è definito come una coppia ordinata (*ordered pair*) di nodi distinti, definita $e_k = \langle v_i, v_j \rangle$, in cui v_i rappresenta l'origine dell'arco (mittente o *sender*) e v_j la destinazione (o *receiver*). È necessario sottolineare la diversa definizione di arco in un grafo diretto rispetto alla definizione di linea in un grafo semplice. La coppia di nodi che definisce un arco nel primo caso è ordinata (per riflettere la direzione del legame tra i due nodi) e nel secondo non è ordinata (codifica solo presenza/assenza del legame). In grafo diretto il numero di archi possibili, dato il numero dei nodi, è dato da $n(n-1)$, poiché per ciascuna coppia di nodi può esistere una coppia di archi orientati in direzioni opposte.

Molti dei concetti applicabili ai grafi semplici possono essere applicati anche a quelli diretti. Alcuni di queste nozioni richiedono una certa riformulazione necessaria per tenere conto delle caratteristiche specifiche degli archi. La nozione di adiacenza, non subisce particolari variazioni, mentre quelle di classi di isomorfismo, diadi, triadi, grado del nodo, percorsi e sentieri devono essere specificati in modo da tenere conto della direzione dell'arco. Anche in un grafo diretto un nodo è incidente con un arco se il nodo è compreso nella coppia ordinata di nodi che definiscono l'arco $e_k = \langle v_i, v_j \rangle$. Nella definizione del concetto di adiacenza è necessario specificare se il nodo nella coppia ordinata è l'originario (o mittente) o è il destinatario. Nel primo caso il nodo v_i è adiacente al nodo v_j se $\langle v_i, v_j \rangle \in \mathcal{E}$, e nel secondo caso il nodo v_j è adiacente dal nodo $\langle v_i, v_j \rangle \in \mathcal{E}$.

Quando un grafo diretto viene rappresentato tramite un digramma i punti rappresentano i nodi e le frecce dirette rappresentano gli archi. L'arco $\langle v_i, v_j \rangle$ è rappresentato da una freccia che parte dal punto che rappresenta v_i e va nel punto che rappresenta v_j .

Grafi segnati, pesati, multipli e ipergrafi. I grafi semplici e diretti possono essere utilizzati per la rappresentazione di relazioni dicotomiche (presenza/assenza di legami) e orientate. Tuttavia, per rappresentare alcuni tipi di

relazioni o caratteristiche delle relazioni più complesse, come i rapporti di simpatia/antipatia o l'intensità del legame tra coppia di attori espressa attraverso valori numerici, è necessaria una generalizzazione del concetto di grafo e di grafo orientato.

Grafi segnati. I grafi segnati (*signed graph*) sono grafi le cui linee sono caratterizzate da un'informazione aggiuntiva, possono assumere segno positivo o negativo, indicati con i segni "+" o "-". Il segno esprime la valenza della relazione: quello positivo rappresenta atteggiamenti o sentimenti positivi (amore, alleanza, ecc.), quello negativo rappresenta atteggiamenti o sentimenti negativi (odio, antagonismo). Il segno è associato ad ogni linea (o arco, in caso di grafo segnato diretto $\mathcal{G}_{d\pm}$) e costituisce un attributo della linea stessa. La definizione di un grafo segnato richiede la formalizzazione di tre set di informazioni: un set di nodi $\mathcal{V} = \{v_1, v_2, \dots, v_n\}$; un set di linee $\mathcal{E} = \{e_1, e_2, \dots, e_l\}$; un set di valenze o segni "+" e "-" associati alle linee $\mathcal{S} = \{s_1, s_2, \dots, s_l\}$. Un grafo segnato viene denominato come $\mathcal{G}_{\pm}(\mathcal{V}, \mathcal{E}, \mathcal{S})$ o \mathcal{G}_{\pm} . Un grafo segnato completo (*complete signed graph*) è un grafo segnato in cui tutte le coppie non ordinate di nodi sono incluse nel set di linee e ad ognuna di esse è associato un segno "+" o "-". In un grafo segnato diretto $\mathcal{G}_{d\pm}$ ogni arco è una coppia ordinata di nodi quindi, l'arco $e_k = \langle v_i, v_j \rangle$ può essere diverso da quello $e_m = \langle v_j, v_i \rangle$ e con esso anche il segno s_k può essere diverso da quello s_m .

Grafi pesati. La rappresentazione di relazioni che esprimono la forza o l'intensità del legame attraverso un valore può avvenire attraverso un grafo pesato o "network" (125). Esempi di relazioni pesate sono la frequenza dell'interazione tra le coppie di persone, l'ammontare degli scambi commerciali tra nazioni, ecc. Per rappresentare tali relazioni è necessario aggiungere un valore o grandezza (*value* o *magnitude*) per ogni linea o arco.

Un grafo pesato o un grafo pesato diretto (*valued graph* o *valued directed graph*) è un grafo (o grafo diretto), in cui ad ogni linea (o arco) è associato un valore. Nella definizione di un grafo pesato è necessaria la formalizzazione di tre set di informazioni: un set di nodi $\mathcal{V} = \{v_1, v_2, \dots, v_n\}$; un set di linee (o archi) $\mathcal{E} = \{e_1, e_2, \dots, e_l\}$; un set di valori (appartenenti al set di numeri reali \mathbb{R}) associati

alle linee (o archi) $\mathcal{W} = \{w_1, w_2, \dots, w_l\}$. Un grafo pesato viene denominato come $\mathcal{G}_{\mathcal{W}} = (\mathcal{V}, \mathcal{E}, \mathcal{W})$ o $\mathcal{G}_{\mathcal{W}}$. Anche nel caso dei grafi pesati valgono le stesse distinzioni effettuate tra grafo semplice e grafo diretto. In particolare, nel grafo pesato semplice vi è un unico valore per la linea tra ogni coppia non ordinata di nodi ($e_k = (v_i, v_j) = (v_j, v_i)$), mentre, nel caso di grafi pesati diretti, gli archi per la coppia ordinata di nodi possono essere differenti ($e_k = \langle v_i, v_j \rangle \neq e_m = \langle v_j, v_i \rangle$), pertanto ad essa vengono associati due valori distinti, w_k e w_m .

Grafi multipli. Viene definito multigrafo (*multigraph* o *multivariate graph*) il grafo in cui è consentita più di una linea (o arco orientato nella stessa direzione, in caso di multigrafo diretto) tra coppie di nodi. Questo tipo di grafo consente di tenere conto di una molteplicità di legami che possono unire gli attori. Un multigrafo \mathcal{G} , $\mathcal{G} = (\mathcal{V}, \mathcal{E}^+)$, è costituito da un set di nodi $\mathcal{V} = \{v_1, v_2, \dots, v_n\}$; una o più set di linee $\mathcal{E}^+ = \{\mathcal{E}_1, \mathcal{E}_2, \dots, \mathcal{E}_R\}$. R rappresenta il numero di set di linee nel multigrafo e il pedice specifica a quale set appartiene la linea. Se ogni relazione non è diretta, ogni linea in ognuna delle R costituisce una coppia non ordinata di nodi distinti, $e_{kr} = (v_i, v_j)$. Una coppia di nodi può essere inclusa in più di un set di linee. Ciascuna coppia non ordinata di nodi può avere da 0 fino a R linee tra loro, tante quante sono i set di R linee.

Ipergrafi. Alcuni network sociali contengono legami tra i sottoinsiemi di attori, come, ad esempio, il legame tra le persone che appartengono allo stesso club o associazione. Gli ipergrafi sono le rappresentazioni appropriate per questo tipo di network, denominati network di affiliazione (*affiliation network* o *membership network*), poiché consentono di specificare subset di nodi, che possono essere di qualsiasi dimensione. Un network di affiliazione è un network *two-mode* costituito da un insieme di attori e una serie di eventi. Ogni evento è un subset di attori di \mathcal{V} . Pertanto, i dati di un network di affiliazione non possono essere pienamente rappresentati in termini di legami a due a due, dal momento che i subset possono includere più di due attori. Un ipergrafo è la rappresentazione appropriata per i dati di un network di affiliazione. Formalmente, un ipergrafo è costituito da un insieme di oggetti e una collezione di subset di oggetti, in cui ogni oggetto appartiene ad almeno un sottoinsieme e nessun sottoinsieme è vuoto. Gli oggetti sono

chiamati punti e gli insiemi di oggetti sono detti legami. In generale, dato il set di punti $\mathcal{A} = \{a_1, a_2, \dots, a_n\}$, e il set dei legami $\mathcal{B} = \{B_1, B_2, \dots, B_m\}$, l'ipergrafo è indicato con $\mathcal{H} = (\mathcal{A}, \mathcal{B})$. Una caratteristica importante di un ipergrafo è che può anche essere descritto dal *dual hypergraph*, denotato \mathcal{H}^* invertendo i ruoli dei punti e legami. In generale, l'ipergrafo $\mathcal{H} = (\mathcal{A}, \mathcal{B})$ ha come set di punti \mathcal{A} e come set di legami \mathcal{B} , quindi il *dual hypergraph* $\mathcal{H}^* = (\mathcal{B}, \mathcal{A})$ ha come set di punti \mathcal{B} e come set di legami \mathcal{A} . Per distinguere tra i punti e i legami, è necessario specificare la notazione. Quando gli elementi \mathcal{A} sono visti come punti e gli elementi in \mathcal{B} sono visti come legami (come nell'ipergrafo $\mathcal{H} = (\mathcal{A}, \mathcal{B})$), si indicano gli elementi \mathcal{A} con lettere minuscole: $\{a_1, a_2, \dots, a_n\}$, e i legami da legami \mathcal{B} usando le lettere maiuscole: $\{B_1, B_2, \dots, B_m\}$. Mentre nel *dual hypergraph*, $\mathcal{H}^* = (\mathcal{B}, \mathcal{A})$ con elementi in \mathcal{B} come punti ed elementi in \mathcal{A} come legami, si indicano gli elementi in \mathcal{A} usando le lettere maiuscole: $\{A_1, A_2, \dots, A_n\}$ e gli elementi in \mathcal{B} con lettere minuscole: $\{b_1, b_2, \dots, b_m\}$.¹ Il concetto di ipergrafo è più generale del concetto di grafo. Il grafo, infatti, può essere definito come un ipergrafo le cui affiliazioni, riguardano solo coppie di nodi e coincidono con ciascun legame tra coppie di soggetti (21, 192).

Sottografi, Diadi, Triadi, Grafi bipartiti. I sottografi sono sottogruppi o subset dei nodi di un grafo e consentono l'analisi di un network sociale a diversi livelli di dettaglio. Per un grafo semplice \mathcal{G} , il grafo \mathcal{G}_s si dice sottografo se il set di nodi di \mathcal{G}_s è un sottoinsieme o subset dei nodi di \mathcal{G} , e se il set di linee di \mathcal{G}_s è un sottoinsieme o subset delle linee di \mathcal{G} . Indicando i nodi e le linee in \mathcal{G}_s rispettivamente con \mathcal{V}_s e \mathcal{E}_s , si può definire come sottografo di \mathcal{G} se $\mathcal{V}_s \subseteq \mathcal{V}$ e se $\mathcal{E}_s \subseteq \mathcal{E}$. Inoltre, in esso tutte le linee in \mathcal{E}_s devono essere definite tra le coppie di nodi in \mathcal{V}_s .

Un sottografo può non includere tutte le linee tra i nodi del sottografo. Ci sono due principali sottografi che possono essere derivati da un grafo semplice a seconda che si considerino un sottoinsieme di nodi e tutte le linee ad essi associati (*node-generated*), oppure, un sottoinsieme di linee e tutti i nodi ad esse incidenti (*line-generated*).

¹La matrice di incidenza associata al *dual hypergraph* è data dalla trasposta della matrice di incidenza associata all'ipergrafo.

Una diade in un grafo semplice è un tipo di sottografo *node-generated*. Essa, infatti, rappresenta una coppia di nodi e le possibili linee tra loro.

Una triade in un grafo semplice è un sottografo costituito da tre nodi e dalle possibili linee tra loro.

Un altro importante tipo di sottografo è generato dalla bipartizione del grafo semplice in due subset di nodi \mathcal{V}_1 e \mathcal{V}_2 tali che ogni linea del grafo è definita da coppie di nodi in cui un nodo appartiene \mathcal{V}_1 e l'altro nodo appartiene \mathcal{V}_2 . In un grafo bipartito (*bipartite graph*) tutte le linee congiungono nodi appartenenti a subset differenti e non esistono linee all'interno dello stesso sottografo, nè ci sono nodi adiacenti nello stesso subset. Si definisce grafo bipartito completo (*complete bipartite graph*), \mathcal{K}_{n_1, n_2} , un grafo in cui ciascun nodo in \mathcal{V}_1 è adiacente a ogni altro in \mathcal{V}_2 , dove n_1 è il numero di elementi in \mathcal{V}_1 , n_2 è il numero di elementi in \mathcal{V}_2 .

Un network *two-mode* con due set di attori e una relazione che lega gli attori in un set agli attori nel secondo set può essere rappresentato come un grafo bipartito. Tuttavia, anche un network *one-mode* può essere rappresentato attraverso un grafo bipartito (ad esempio, se la relazione è asimmetrica)(124, 192) .

Il partizionamento dei nodi in un grafo può essere generalizzato da due a più subset $\mathcal{V}_1, \mathcal{V}_2, \dots, \mathcal{V}_s$. Un grafo *s-partito* (*s-partite graph*), è un grafo in cui i nodi sono partizionati in s subset e le linee congiungono esclusivamente nodi appartenenti a subset diversi. Anche la nozione di grafo bipartito completo può essere estesa ad un grafo *s-partito* completo (*complete s-partite graph*). Un grafo è *s-partito* completo se tutte le possibili coppie di nodi appartenenti a subset diversi sono adiacenti, cioè sono presenti tutte le linee tra subset e non ci sono linee incidenti con due nodi appartenenti allo stesso subset. In un grafo diretto una diade è un sottografo costituito da due nodi e da tutti i possibili archi tra essi. In un grafo diretto una diade può assumere tre classi di isomorfismi: diade nulla, non è presente nessun arco (né $\langle v_i, v_j \rangle$ né $\langle v_j, v_i \rangle$); diade asimmetrica, è presente un solo arco in una delle due direzioni ($\langle v_i, v_j \rangle$ o $\langle v_j, v_i \rangle$); diade simmetrica o reciproca, sono presenti gli archi in entrambe le direzioni ($\langle v_i, v_j \rangle$ e $\langle v_j, v_i \rangle$).

I tipi di diadi in un grafo orientato sono ampiamente utilizzati per descrivere un network sociale. La propensione alla reciprocità o all'asimmetria possono essere

sintetizzate contando il numero di diadi per ciascuna delle tre classi di isomorfismo. Molto utilizzati sono anche i sottografi di dimensione tre, le triadi, coinvolti nello studio di concetti centrali nella SNA (*balance*, *clusterability*, transittività).

In un grafo segnato una diade può assumere tre forme distinte: linea positiva, linea negativa, assenza di linea. In un grafo segnato diretto le forme possibili diventano nove, a seconda di tutte le possibili combinazioni di segni in relazioni simmetriche, asimmetriche e in caso di assenza di legame.

Per l'analisi delle triadi in un grafo segnato è necessario ricorrere al concetto di ciclo definito come una sequenza alternata di nodi e linee distinte, escluso il nodo iniziale, che coincide con quello finale della sequenza. In un grafo segnato è possibile attribuire un segno a qualsiasi ciclo, moltiplicando tra loro i segni delle linee comprese nel ciclo stesso. Il prodotto dei segni sarà positivo se tutti i segni della sequenza sono positivi o se nella sequenza esiste un numero pari di segni negativi; altrimenti sarà negativo. Questa regola è stata utilizzata da (125) nell'ambito della teoria dell'equilibrio strutturale nei grafi segnati, secondo la quale un grafo segnato è in equilibrio se e solo se tutti i suoi cicli sono positivi trasportatori (62, 192).

1.3.1.2 Proprietà dei nodi e proprietà dei grafi

Degree. In un grafo semplice il grado (*degree*) di un nodo $d(v_i)$, è dato dal numero di linee incidenti con esso o dal numero di nodi adiacenti ad esso, noti anche come vicini (*neighbour*). Il *degree* assume valori da 0 (nessun nodo adiacente ad un certo nodo, caso in cui il nodo è isolato) ad un massimo di $n - 1$ (il nodo è adiacente a tutti gli altri). Il grado di un nodo rappresenta una prima indicazione sulle caratteristiche strutturali degli attori; esso è una misura del grado di attività del nodo nella rete (192). Per avere una misura sintetica del *degree* di tutti gli attori del grafo e della sua variabilità, è possibile calcolare il grado medio dei nodi di un grafo e la relativa varianza. Ciò consente di avere informazioni sul livello e le differenze di integrazione tra gli attori del network (62).

Il *degree* medio è dato da:

$$\bar{d} = \frac{\sum_{i=1}^n d(v_i)}{n} = \frac{2l}{n}$$

, la varianza dei *degree* è data da:

$$S_D^2 = \frac{\sum_{i=1}^n (d(v_i) - \bar{d})^2}{n}$$

Un grafo è *d-regular* (dove d è un valore costante per tutti i *degree*) se i *degree* di tutti i nodi sono uguali, tale misura esprime il grado di uniformità. Quando un grafo è *d-regular* la sua varianza è nulla $S_D^2 = 0$.

La variabilità nei *degree* dei nodi restituisce un'informazione sui differenti gradi di attività degli stessi nel network, misurata a partire dal numero di legami che hanno con gli altri nodi.

Il concetto di adiacenza riadattato per grafi diretti, consente di ridefinire anche altre proprietà strutturali come il grado (*degree*) di un nodo. In un grafo diretto, un nodo può essere sia adiacente "a", sia adiacente "da" un altro nodo, a seconda della direzione dell'arco. Si possono distinguere l'*indegree* (grado in entrata) $d_I(v_i)$ e l'*outdegree* (grado in uscita) $d_O(v_i)$. L'*indegree* è uguale al numero di nodi che sono adiacenti ad v_i ossia il numero di archi che terminano in v_i . L'*outdegree* è uguale al numero di nodi che sono adiacenti da v_i ossia il numero di archi che hanno origine dal nodo v_i . Da un punto di vista applicativo i concetti di *indegree* e di *outdegree* possono essere considerati misure della espansività, il primo, e della di ricettività o di popolarità, il secondo. In una relazione di amicizia, un attore con una elevato *indegree* è colui che viene indicato da molti altri attori come amico, quindi, ha un elevato grado di apprezzamento. Un attore con un elevato *outdegree* è colui che indica molti attori come amici, mostrando una certa disposizione verso gli altri e adesione al gruppo (62, 192).

Come per il *degree* è utile sintetizzare l'*indegree* e l'*outdegree* attraverso il loro valore medio.

L'*indegree* medio \bar{d}_I è dato da:

$$\bar{d}_I = \frac{\sum_{i=1}^n d_I(v_i)}{n}$$

L'*outdegree* medio \bar{d}_O è dato da:

$$\bar{d}_O = \frac{\sum_{i=1}^n d_O(v_i)}{n}$$

Dato che $\sum_{i=1}^n d_I(v_i) = \sum_{i=1}^n d_O(v_i) = l$, ne consegue che

$$\bar{d}_I = \bar{d}_O = \frac{l}{n}$$

Un'informazione di un certo interesse è legata alla variabilità dell'*indegree* e dell'*outdegree*.

La varianza dell'*indegree* $S_{D_I}^2$ è data da:

$$S_{D_I}^2 = \frac{\sum_{i=1}^n (d_I(v_i) - \bar{d}_I)^2}{n}$$

e la varianza dell'*outdegree* $S_{D_O}^2$ è data da:

$$S_{D_O}^2 = \frac{\sum_{i=1}^n (d_O(v_i) - \bar{d}_O)^2}{n}$$

Entrambi gli indici restituiscono una misura di quanto gli attori nel network si differenzino in termini di invio e ricezione di legami, quindi forniscono una misura della centralizzazione del network.

Nei grafi diretti, la nozione di arco in entrata e in uscita, di *indegree* e *outdegree*, consente di distinguere diversi tipi di nodi, a seconda della presenza e della direzione degli archi: nodi in cui $d_I(v_i) = d_O(v_i) = 0$ definiti come nodi isolati; nodi con $d_I(v_i) = 0$ e $d_O(v_i) > 0$, definiti come trasmettitori; nodi con $d_I(v_i) > 0$ e $d_O(v_i) = 0$, definiti riceventi; nodi con $d_I(v_i) > 0$ e/o $d_O(v_i) > 0$ chiamati nodi ordinari; nodi con $d_I(v_i) = d_O(v_i) = 1$, chiamati trasportatori (62, 192).

Nel caso di grafi pesati la presenza di valori diversi associati alle linee (o gli archi in caso di grafi pesati diretti) rende più complicato il calcolo di alcune misure strutturali. Ad esempio il grado di un nodo dove tener conto non solo del numero delle linee incidenti con esso, ma anche del loro valore. A tal proposito è possibile calcolare la media dei valori delle linee incidenti o la loro somma. In un grafo pesato la lunghezza di un sentiero o un percorso è dato dalla somma dei pesi (delle linee che ne fanno parte).

Densità. Un'altra rilevante misura delle caratteristiche strutturali del network è legata alla proporzione di linee presenti nell'intero grafo.

Il numero massimo di legami possibili in un grafo dipende del numero dei nodi. Considerato n come numero di nodi presenti nel grafo, il numero di linee possibili

nel caso dei un grafo semplice sono $n(n-1)/2$. In un grafo semplice la densità (*density*) è data dalla proporzione tra i legami effettivamente presenti e quelli possibili, dato il numero di nodi:

$$\Delta = \frac{l}{n(n-1)/2} = \frac{2l}{n(n-1)}$$

L'indice di densità assume valori che vanno da 0, in caso di assenza di linee ($l = 0$), a 1 nel caso in cui sono presenti tutte le linee possibili ($l = n(n-1)/2$). Quando sono presenti tutte le linee, tutti i nodi sono adiacenti (il grado di ciascun nodo è $n-1$), il grafo viene detto completo (*complete*). Seguendo la stessa logica è possibile calcolare la densità anche per i sottografi per misurarne il grado di coesione.

La densità di un grafo diretto è data dalla proporzione di archi presenti. Dato che un arco è una coppia ordinata di nodi, possono esserci $n(n-1)$ possibili archi:

$$\Delta = \frac{l}{n(n-1)}$$

La densità di un grafo pesato, o di un grafo diretto pesato è definita come il rapporto tra il numero di linee o archi presenti per il massimo possibile. La densità può essere vista anche come la media dei valori assegnati alle linee/archi. Viene assegnato il valore 1 per la presenza di ogni linea o arco, un valore 0 per l'assenza. La somma di questi valori è uguale al numero di linee o archi; questa somma può essere divisa per il suo massimo valore possibile. Generalizzando il concetto di densità per un grafo pesato o per un grafo diretto, si può calcolare la media dei valori associati alle linee/archi in tutte le linee/archi. In questo modo, per un grafo pesato o per un grafo diretto, la densità è data da:

$$\Delta = \frac{\sum w_k}{n(n-1)}$$

in cui la somma è effettuata per ogni k linee del grafo. Questa misura la forza media delle linee o degli archi nel grafo pesato o nel grafo diretto. Si deve considerare che in questo caso l'indice può assumere valori superiori ad 1, anche se non tutte le coppie di nodi sono tra loro adiacenti.

Percorsi, Sentieri, Cicli. I concetti di adiacenza, incidenza e densità sono basati sull'analisi delle relazioni dirette tra coppie di nodi. Tuttavia è possibile analizzare alcune proprietà del network considerando il modo in cui i nodi sono legati attraverso legami indiretti. Queste proprietà sono utili per analizzare il grado di connessione del grafo, per definire il concetto di distanza tra coppie di nodi, e per individuare nodi e linee rilevanti per la connettività del grafo.

In un grafo semplice, un percorso (*walk*), \mathcal{W} , è una sequenza di nodi e linee che inizia con un nodo e finisce con un nodo, in cui ogni nodo è incidente con la linea che lo precede e con quella che lo segue nella sequenza. La sequenza di un percorso, in parte, può contenere più di una volta gli stessi nodi e le stesse linee, e il nodo di partenza può essere diverso o può coincidere con quello di arrivo. In un grafo semplice, un percorso può essere indicato enumerando in ordine di sequenza i nodi e le linee, partendo dal nodo di origine e terminando con quello di destinazione. La lunghezza di un percorso è data dal numero di linee enumerate nella sequenza. L'inverso di un percorso, indicato con \mathcal{W}^{-1} , è dato dalla sequenza del percorso elencato nell'ordine opposto, iniziando dal nodo di destinazione e terminando con quello di origine. Se il nodo di partenza corrisponde a quello di destinazione il percorso si dice chiuso (*closed*). Tipi particolari di percorsi sono la traccia (*trail*), il sentiero (*path*) e il ciclo (*cycle*).

Una traccia (*trail*) è un percorso in cui le linee possono essere inserite solo una volta. Un sentiero (*path*), \mathcal{P} , è un percorso in cui nodi e linee possono essere inseriti solo una sola volta. Se tra due nodi v_i e v_j esiste un sentiero, essi vengono definiti raggiungibili (*reachable*).

Percorso, traccia e sentiero costituiscono possibili passaggi di generalità decrescente attraverso le coppie di nodi. In un grafo, in genere, il sentiero è l'attraversamento più breve possibile. Un ciclo (*cycle*) è un sentiero chiuso costituito da almeno tre nodi in cui tutte le linee sono incluse solo una volta, e tutti i nodi ad eccezione del nodo di inizio e destinazione sono incluse solo una volta. Un grafo che non contiene cicli è detto aciclico (*acyclic*). I concetti di percorso e sentiero nel caso dei grafi diretti devono essere riformulati per tenere conto della direzione degli archi. Un percorso diretto (*directed walk*) è dato da una sequenza alternata di nodi e archi tale che ogni arco ha origine nel nodo precedente e destinazione nel nodo successivo. Un sentiero diretto (*direct path*) è dato da un percorso diretto

in cui nessun arco e nessun nodo vengono inclusi più di una volta. La lunghezza di un sentiero, come nel caso del grafo semplice, è data dal numero di archi della sequenza. Un ciclo diretto (*directed cycle*) è un sentiero diretto in cui la sequenza inizia e finisce con lo stesso nodo.

Allentando il vincolo della sequenza unidirezionale, diviene possibile enumerare una qualsiasi sequenza senza tenere conto della direzione degli archi. In questo caso tali concetti prendono il nome di semi-percorso (*semiwalk*), semi-sentiero (*semipath*), semi-ciclo (*semicycle*).

Grafi connessi e Componenti. Una proprietà importante di un grafo è legata alla sua connessione. Un grafo semplice si dice connesso (*connected graph*) se esiste un sentiero (*path*) tra ogni coppia di nodi. Se un grafo è connesso, tutte le coppie di nodi sono raggiungibili. Se un grafo non è connesso si dice sconnesso (*disconnected graph*).

All'interno di un grafo disconnesso, \mathcal{G} , i nodi possono essere partizionati in subset che non sono legati tra loro tramite sentieri (*path*). Tali subset costituiscono dei sottografi a massima connessione, \mathcal{G}_s (ognuno costituisce un massimo sottografo connesso) dette componenti (*component*) se tutti i nodi al loro interno, \mathcal{V}_s , sono tra loro connessi e l'inclusione di un ulteriore nodo provoca la disconnessione di \mathcal{G}_s .

In una componente tutte le coppie di nodi del sottografo sono raggiungibili, ossia esiste un sentiero tra loro. Ne consegue che il grafo si dice connesso se vi è un solo componente, per contro, si dice sconnesso se ve ne sono più di uno. Dato un grafo \mathcal{G} connesso o disconnesso, è possibile individuare uno o più sottografi \mathcal{G}_s , definiti *clique*, se la densità interna è uguale a uno, cioè se tutte le coppie di nodi in \mathcal{V}_s sono tra loro direttamente collegate da una linea. Una *clique* viene definita come il massimo sottografo completo.

Distanza geodetica, Eccentricità, Diametro. Il concetto di sentiero (*path*) è alla base della definizione del concetto di distanza nel grafo. La distanza tra due nodi, o distanza geodetica (*geodesic distance*), è definita come la lunghezza di una geodetica tra loro. Quest'ultima è data dal sentiero più breve tra due nodi. Il sentiero più breve tra v_i e v_j , o distanza, viene indicato con $d(i, j)$. Se

non vi è alcun percorso tra i due nodi (ossia, non sono raggiungibili), allora la distanza tra di loro è infinita (o non definita). Se un grafo non è connesso, allora la distanza tra almeno una coppia di nodi è infinito (poiché la distanza tra due nodi tra componenti differenti è infinita).

In un grafo semplice, una geodetica tra v_i e v_j è uguale alla geodetica tra v_j e v_i . Di conseguenza, la distanza tra v_i e v_j è uguale alla distanza tra v_j e v_i , pertanto $d(i, j) = d(j, i)$. La distanza più lunga tra un nodo e qualsiasi altro viene detta eccentricità (*eccentricity*) del nodo. In termini formali l'eccentricità di un nodo v_i è uguale alla massima $d(i, j)$ per ogni v_j , quindi $\max_j d(i, j)$. Si noti che la distanza può essere calcolata per ogni coppia di nodi, mentre l'eccentricità è riferita ad un singolo nodo, e implica il confronto delle distanze di quel nodo rispetto a tutti gli altri del grafo. Infine, si indica con il termine diametro (*diameter*) di un grafo, la geodetica più lunga, ossia l'eccentricità più lunga, misurabile per un qualsiasi nodo del grafo stesso. Il diametro di un grafo è uguale alla massima $d(i, j)$ per ogni v_i e v_j , quindi $\max_i \max_j d(i, j)$, e può assumere valori tra 1, se il grafo è completo, e $n - 1$, se tutti i nodi sono tra loro connessi da linee che formano un unico sentiero. Il concetto di diametro di un grafo è importante poiché quantifica la distanza tra i due nodi che sono agli estremi del grafo.

A differenza del concetto di densità, il cui calcolo può essere effettuato su qualsiasi tipo di grafo, in quanto necessita solo del numero dei nodi e quello dei legami, indipendentemente dalla loro collocazione nel grafo stesso, i concetti di percorso, sentiero, distanza, eccentricità e diametro possono essere applicati solo ad insiemi di nodi tra loro connessi cioè collegati direttamente o indirettamente da almeno un percorso(62) .

In un grafo diretto, il sentiero tra i nodi v_i a v_j può essere diverso da quello tra i nodi v_j e v_i essendoci il vincolo dell'orientamento di tutti gli archi nella stessa direzione. Pertanto, la distanza nella coppia ordinata di nodi v_i e v_j può differire dalla distanza tra i nodi v_j e v_i . Ciò implica che in un grafo debolmente connesso possono esserci coppie ordinate di nodi non raggiungibili, quindi a distanza indefinita (o infinita). Per quanto riguarda il diametro di grafi diretti, come nel caso dei grafi semplici, esso è definito dalla lunghezza della distanza più lunga. Tuttavia, nei grafi diretti, è necessaria l'esistenza di un sentiero tra un nodo e ogni altro nodo del grafo. Tale vincolo rende calcolabile il diametro soltanto in

caso di connessione forte o ricorsiva. Per altre forme di connessione la distanza è indefinita (o infinita).

Misure di connessione. Grazie alla combinazione dei concetti di raggiungibilità tra coppie di punti, di grafo connesso e di componenti in un grafo disconnesso, è possibile definire alcuni concetti importanti per individuare punti e linee critiche per la connettività del grafo. Essa ha a che fare con la possibilità che un grafo rimanga connesso anche quando i nodi e/o le linee vengono eliminate. Di conseguenza, sono concetti che consentono di misurare il livello di connessione del grafo (*connectivity*).

Un punto di separazione (*cutpoint*) è un nodo appartenente ad un componente, tale che la sua eliminazione comporta la separazione del componente stesso in almeno due componenti distinti.

Un nodo, v_i , è un punto di separazione se il numero di componenti nel grafo che contiene il nodo è inferiore rispetto al numero di componenti del sottografo che si otterrebbe rimuovendo v_i dal grafo. Anche un subset di nodi (*cutset*), di dimensione k , può costituire un punto di separazione se occupa una posizione nella struttura del grafo tale per cui la loro rimozione genera una divisione del componente originario in almeno due componenti. Un concetto simile a quello di punto di separazione, utilizzato per i nodi, può essere applicato anche per le linee. Si definisce ponte (*bridge*) quella linea, appartenente ad un grafo connesso \mathcal{G} , la cui rimozione comporta una divisione del grafo originario in un numero di componenti maggiori rispetto al numero di componenti che si otterrebbero senza la rimozione della linea. Come nel caso del set di nodi anche per le linee è possibile individuare un subset di linee l — *linecut* la cui rimozione genera la disconnessione del grafo, ossia un numero di componenti maggiori rispetto a quelle presenti nel grafo originario. Ad un livello sostanziale/applicativo, i due concetti di punto di separazione e ponte vengono utilizzati nell'analisi della connessione dei network e nell'individuazione di soggetti o tipi di legami che grazie alla loro posizione strutturale influenzano l'integrazione del gruppo o possono svolgere importanti funzioni di mediazione tra gruppi diversi (62) .

Considerando che la misura del livello di connessione in un grafo tiene conto del numero di nodi, o il numero di linee, che devono essere rimossi per lasciare il grafo

disconnesso, essa può essere vista come misura della coesione e della robustezza del grafo (192).

Il grado di coesione di un grafo può essere valutato a partire dalla numerosità delle linee, dal livello del *degree* dei nodi, e dalla lunghezza e numerosità dei sentieri tra coppie di nodi. Pertanto, un grafo coeso è caratterizzato da numerose geodetiche brevi e piccoli diametri. Se un grafo non è coeso, allora è vulnerabile rispetto alla rimozione di alcuni nodi o linee.

A partire dalle nozioni di *cutset* e di *line cut* è possibile definire due misure del livello di connessione di un grafo semplice, una relativa alla rimozione di nodi, e l'altra basata sulla rimozione di linee(192, 199) .

Il punto di connessione (*point-connectivity*) o nodo di connessione (*node-connectivity*) di un grafo, $\kappa(\mathcal{G})$, è il numero minimo k per il quale il grafo ha un k -nodi di separazione. Quindi, il punto di connessione è il numero minimo di nodi che devono essere rimossi per rendere il grafo disconnesso. Se $k = 0$ il grafo è disconnesso poiché nessun nodo deve essere rimosso. Se $k = 1$ il grafo contiene un punto di separazione, poiché la rimozione del nodo lascia il grafo disconnesso. Se $k = 2$ vuol dire che il grafo risulta disconnesso se contemporaneamente viene rimossa una coppia di nodi, quindi due è il numero minimo di nodi che deve essere rimosso per rendere il grafo disconnesso. Di conseguenza, valori elevati di k indicano livelli elevati di connessione e per valori minori di k il grafo si dice *k-node connected*. La linea di connessione (*line-connectivity*) o legame di connessione (*edge-connectivity*) di un grafo, $\lambda(\mathcal{G})$, è il numero minimo λ per il quale il grafo ha un λ -*line cut*. Il valore λ corrisponde al numero minimo di linee che devono essere rimosse per disconnettere il grafo. Se $\lambda(\mathcal{G} \geq l)$, il grafo si dice *l-line connected*.

Maggiori sono i nodi di connessione, o le linee di connessione, di un grafo e minore è la vulnerabilità dello stesso, ossia il rischio di diventare disconnesso.

Considerando le varie nozioni strutturali riadattate al caso dei grafi diretti diviene possibile specificare anche il concetto di connessione appropriato per tale tipo di grafo. Se esiste un sentiero diretto dal nodo v_i e v_j significa che il nodo v_j è raggiungibile dal nodo v_i . In un grafo diretto, possono essere distinti quattro tipi di connessione tra coppie di nodi a seconda che si considerino i sentieri o i semi-sentieri, ossia si consideri o meno la direzione dell'arco.

Una coppia di nodi v_i, v_j può essere: (i) debolmente connessa (*Weakly connected*) se è unita da un semi-sentiero; (ii) connessa unilateralmente (*Unilaterally connected*) se è unita da un sentiero da v_i a v_j , o un sentiero da v_j a v_i ; (iii) fortemente connessa (*Strongly connected*) se esiste un sentiero da v_i a v_j , e un sentiero da v_j a v_i (i due sentieri possono contenere anche nodi e archi diversi); (iv) ricorsivamente connessa (*Recursively connected*) se è fortemente connessa, e il sentiero v_i a v_j contiene gli stessi nodi e archi del sentiero da v_j a v_i , in ordine inverso (192). Ognuno di essi esprime una definizione di connessione via via più stringente/rigida. Questi tipi di connessione tra coppie di nodi danno origine a quattro possibili tipi di connessione per il grafo diretto nel suo complesso. Un grafo diretto può essere: (i) debolmente connesso (*Weakly connected*) se tutte le coppie di nodi sono debolmente connesse; (ii) connesso unilateralmente (*Unilaterally connected*) se tutte le coppie di nodi sono connesse unilateralmente; (iii) fortemente connesso (*Strongly connected*) se tutte le coppie di nodi sono fortemente connesse; (iv) ricorsivamente connesso (*Recursively connected*) se tutte le coppie di nodi sono ricorsivamente connesse.

Grafi isomorfi. Un importante nozione è legata al confronto tra grafi ed è quella di isomorfismo (*isomorphism*). Gli isomorfismi tra grafi sono di notevole interesse poiché se due grafi sono isomorfi, allora sono identici rispetto a tutte le proprietà teoriche del grafo (192). Due grafi, \mathcal{G} e \mathcal{G}^* sono isomorfi se esiste una corrispondenza "uno a uno" tra i nodi di \mathcal{G} e i nodi di \mathcal{G}^* che consenta di preservare l'adiacenza dei nodi. Una corrispondenza biunivoca implica che ad ogni nodo di \mathcal{G} corrisponda uno, ed uno solo, nodo di \mathcal{G}^* e viceversa.

Siano $\mathcal{V} = \{v_1, v_2, \dots, v_n\}$ i nodi in \mathcal{G} e siano $\mathcal{V}^* = \{v^*_1, v^*_2, \dots, v^*_n\}$ i nodi in \mathcal{G}^* . La notazione per indicare che al nodo v_i in \mathcal{G} corrisponde il nodo v^*_k in \mathcal{G}^* è data da $\phi(v_i) = v^*_k$. L'inverso di tale corrispondenza ϕ^{-1} è quella che associa il nodo v^*_k in \mathcal{G}^* al nodo v_i in \mathcal{G} ossia $\phi^{-1}(v^*_k) = v_i$. Dato che la corrispondenza è biunivoca vale $\phi(n_i) = n^*_k$ se e solo se $\phi^{-1}(v^*_k) = v_i$. Formalmente due grafi sono isomorfi se per tutti i nodi $v_i, v_j \in \mathcal{V}$ e quelli $v^*_k, v^*_l \in \mathcal{V}^*$ esiste una corrispondenza biunivoca, $\phi(v_i) = v^*_k$ e $\phi(v_j) = v^*_l$ tale che nodi $l_m = (v_i, v_j) \in \mathcal{E}$ se e solo se $l_o = (v^*_k, v^*_l) \in \mathcal{E}^*$. Due nodi adiacenti in un grafo devono essere associati in modo che siano adiacenti anche nel grafo isomorfo. Due grafi isomorfi

hanno le stesse proprietà strutturali, ad esempio hanno lo stesso numero di nodi, linee, diametro, ecc. Di conseguenza, le proprietà teoriche che valgono per il grafo \mathcal{G} valgono anche per il suo isomorfo \mathcal{G}^* . Le caratteristiche associate al grafo nel suo complesso valgono anche per i suoi nodi. Infatti, se due grafi in \mathcal{G} e \mathcal{G}^* sono isomorfi, e ad v_i nel grafo \mathcal{G} è associato il nodo v_k^* in \mathcal{G}^* , ossia $(\phi(v_i) = v_k^*)$, allora v_i ed v_k^* sono identici in tutte le loro proprietà strutturali (stesso *degree* del nodo, eccentricità, ecc.). Questa proprietà è molto importante per la definizione di alcuni tipi di equivalenze posizionali (192).

Allo stesso modo di nodi isomorfi è possibile definire sottografi isomorfi. Due sottografi, \mathcal{G} e \mathcal{G}^* sono isomorfi se esiste una corrispondenza biunivoca tra i nodi di \mathcal{G} e i nodi di \mathcal{G}^* che consenta di preservare l'adiacenza dei nodi. Sottografi isomorfi appartengono alla stessa classe di isomorfismo (*isomorphism class*).

Grafi speciali. Di una certa utilità a livello concettuale ma anche operativo è la definizione di alcuni particolari tipi di grafo dotati di specifiche proprietà.

Il complemento (*complement*), $\bar{\mathcal{G}}$, di un grafo semplice \mathcal{G} , ha lo stesso insieme di nodi del grafo \mathcal{G} . In $\bar{\mathcal{G}}$ è presente una linea tra una coppia di nodi solo se la coppia di nodi non è presente nel set di linee in \mathcal{G} , e, viceversa, una linea non è presente in $\bar{\mathcal{G}}$ se è presente in \mathcal{G} . Pertanto, se i nodi v_i e v_j sono adiacenti in \mathcal{G} non saranno adiacenti in $\bar{\mathcal{G}}$, e viceversa.

Un grafo che è connesso ed è aciclico viene detto albero (*tree*). I grafi ad albero contengono il numero minimo di linee per essere connessi e non contengono cicli. Essi possiedono almeno tre importanti caratteristiche. Sono minimamente connessi, poiché ogni linea è un ponte (*line cut*). Il numero di linee in un grafo ad albero è dato dal numero di nodi meno uno ($l = n - 1$). In un grafo ad albero tra ogni due nodi esiste solo un sentiero.

Un grafo disconnesso che non contiene cicli è detto foresta (*forest*), in esso ogni componente è un albero. Il numero di linee in un albero o in una foresta è uguale al numero di nodi meno il numero di componenti del grafo. In un albero il numero di linee è $l = n - 1$ perché il numero di componenti è 1.

In un grafo diretto possono essere individuati due tipi di grafi speciali. Questi grafi diretti possono essere utilizzati per rappresentare l'opposto e la negazione di una relazione.

Il complemento (*complement*), $\bar{\mathcal{G}}_d$, di un grafo diretto \mathcal{G}_d , ha lo stesso insieme di nodi del grafo \mathcal{G}_d . In $\bar{\mathcal{G}}_d$ è presente un arco tra una coppia ordinata di nodi solo se la coppia di nodi non è presente nel set di archi in \mathcal{G}_d , e, viceversa, un arco non è presente in $\bar{\mathcal{G}}_d$ se è presente in \mathcal{G}_d . Pertanto, se l'arco $\langle v_i, v_j \rangle$ è presente in \mathcal{G}_d non sarà presente in $\bar{\mathcal{G}}_d$, e viceversa.

L'inverso (*converse*), \mathcal{G}'_d , di un grafo diretto \mathcal{G}_d , ha lo stesso insieme di nodi del grafo \mathcal{G}_d ; tuttavia l'arco $\langle v_i, v_j \rangle$ è presente in \mathcal{G}'_d solo se l'arco $\langle v_j, v_i \rangle$ è presente in \mathcal{G}_d . L'inverso è ottenuto invertendo la direzione di tutti gli archi del grafo diretto originario. Di conseguenza, gli archi del grafo inverso collegano le stesse coppie di nodi del grafo diretto originario, ma tutti gli archi sono in direzione inversa. L'inverso di un grafo orientato può essere utile per concettualizzare relazioni che sottendono delle opposizioni (relazioni di dominio/sottomissione). Il complemento di un grafo orientato, invece, può essere utile per concettualizzare l'assenza di un legame, o *non* relazione (amico scelto/non scelto).

1.3.2 Matrici di dati relazionali: matrici di Adiacenza e di Incidenza

Le informazioni contenute in un grafo \mathcal{G} possono essere organizzate anche in forma matriciale. Sebbene la rappresentazione tramite grafo sia caratterizzata dall'immediatezza delle intuizioni visive, tale strumento diventa meno informativo con l'aumentare delle dimensioni del grafo e della quantità e complessità dei dati relazionali. L'organizzazione in matrice consente l'analisi delle proprietà strutturali del grafo sfruttando i vantaggi e le proprietà derivanti dall'applicazione delle operazioni algebriche. Infatti, molte caratteristiche fondamentali dei network sociali possono essere analizzate attraverso la manipolazione diretta di matrici (trasposizione, addizione, moltiplicazione di matrici)¹, tutte operazioni che permettono di estrarre informazioni sulla loro struttura (176). Due sono le matrici più utilizzate nella SNA: la matrice di adiacenza (o sociomatrice) e quella di incidenza. Tali matrici contengono tutte le informazioni presenti in un grafo. Quindi, l'insieme

¹Per una trattazione dettagliata sulle operazioni per matrici derivate da grafi si rimanda a Wasserman e Faust (192), Wilson (199).

di nodi e linee in un grafo è completamente descritto dalle informazioni di una matrice (124, 192, 199).

Matrici di adiacenza. La matrice più utilizzata nella SNA è quella di adiacenza (*adjacency matrix* o *sociomatrix*), la quale prende il nome dalle informazioni che contiene, infatti codifica l'adiacenza, o meno, di due nodi. Una matrice di adiacenza \mathbf{X} di dimensioni $n \times n$ (n righe e n colonne) rappresenta un network *one-mode*. Essa presenta una riga e una colonna per ogni nodo, con righe e colonne etichettate da $1, 2, \dots, n$. Le righe e le colonne indicano i nodi del grafo, o gli attori del network, nello stesso ordine. x_{ij} sono le entrate di una matrice di adiacenza e indicano se la coppia di nodi è adiacente. $x_{ij} = 1$ se nella (i, j) –esima cella (riga i , colonna j) vi è una linea tra v_i e v_j e $x_{ij} = 0$ se la linea non esiste. I valori binari (*binary data*) codificano la presenza/assenza della relazione, legame alla base di un grafo semplice in cui le linee non presentano direzione, valenza e valori (pesi). Se tra v_i e v_j esiste una linea essa esiste anche da v_j a v_i ; pertanto, $x_{ij} = 1$, e $x_{ji} = 1$. Pertanto, una matrice di adiacenza associata ad un grafo semplice è simmetrica. Una matrice è simmetrica se $x_{ij} = x_{ji}$ per tutti i e j ; ne consegue che i valori delle entrate nei triangoli in alto a destra e in basso a sinistra sono identici. Le entrate sulla diagonale, corrispondenti a x_{ii} , non sono definite, se nel grafo non sono ammessi *loop*. Una matrice di adiacenza per un grafo completo contiene 1 in tutte le celle extra-diagonali. Dal momento che tutti i nodi sono adiacenti, $x_{ij} = x_{ji} = 1$ per tutti $i \neq j$. Una matrice di adiacenza per un grafo vuoto contiene 0 ovunque, dato che non ci sono nodi adiacenti(192).

La matrice di adiacenza di un grafo diretto ha entrate $x_{ij} = 1$ se vi è un arco $\langle v_i, v_j \rangle$ in E , ossia se vi è un arco dalla riga che corrisponde al nodo v_i alla colonna che corrisponde al nodo v_j , e $x_{ij} = 0$ altrimenti. L'entrata x_{ji} potrebbe differire da x_{ij} indicando una "scelta" non reciproca.

Anche un grafo pesato, può essere rappresentato attraverso una matrice di adiacenza. In questo caso l'entrata in x_{ij} è data dal valore, o peso, associato alla linea tra il nodo v_i e il nodo v_j in un grafo pesato semplice, o dal valore associato all'arco da v_i a v_j in un grafo pesato diretto. Nel primo caso l'entrata x_{ij} è un unico valore, w_k , associato alla linea $e_k = (v_i, v_j)$, e quindi $x_{ij} = x_{ji} = w_{ij}$. Nel secondo caso, per un grafo pesato diretto, l'arco $e_k = \langle v_i, v_j \rangle$, con il valore w_k ,

e l'arco $e_m = \langle v_j, v_i \rangle$, con il valore w_m possono differire, pertanto, in matrice, sono presenti due valori distinti, $x_{ij} = w_k$ e $x_{ji} = w_m$. L'entrata nella cella (i, j) di \mathbf{X} rappresenta la forza o l'intensità del legame da attore i all'attore j .

Matrici di incidenza. La seconda matrice che può essere utilizzata per rappresentare le informazioni contenute in un grafo è la matrice di incidenza (*incidence matrix*), \mathbf{F} o $\mathbf{F}(\mathcal{G})$, che registra quali linee sono incidenti con quali nodi. \mathbf{F} è una matrice di dimensioni $n \times l$, con n nodi e l linee; una riga per ogni nodo, e una colonna per ogni linea. L'entrata $f_{ij} = 1$ se il nodo v_i è incidente con la linea e_j ed $I_{ij} = 0$ se il nodo v_i non è incidente con la linea e_j . Poiché la linea $e_k = (v_i, v_j)$ è incidente con i due nodi v_i e v_j , ogni colonna in \mathbf{F} ha esattamente due 1 in esso, registrando l'incidenza di due nodi con la linea. La matrice di incidenza è binaria, poiché una linea o è incidente con un nodo o non lo è, ma non è necessariamente quadrata.

Una matrice di incidenza può essere utilizzata per riprodurre le informazioni contenute sia in un network *two-mode* diadico sia in un network *two-mode* di affiliazione. Un network *two-mode* diadico $\mathcal{N}(\mathcal{V}_1, \mathcal{V}_2, R)$ con due diversi set di attori, \mathcal{V}_1 e \mathcal{V}_2 , e una relazione, R , che lega gli attori del primo set agli attori del secondo set può essere rappresentato attraverso un grafo bipartito $\mathcal{G}(\mathcal{V}_1, \mathcal{V}_2, \mathcal{E})$, con $\mathcal{V}_1 \cup \mathcal{V}_2 = \mathcal{V}$ e $\mathcal{V}_1 \cap \mathcal{V}_2 = \emptyset$, le cui informazioni possono essere registrate in una matrice di incidenza \mathbf{F} di dimensioni $n \times m$, dove le righe rappresentano i nodi (attori) del primo set $\mathcal{V}_1 = \{a_1, a_2, \dots, a_n\}$ e le colonne rappresentano i nodi del secondo set (attori con cui essi sono legati) in $\mathcal{V}_2 = \{e_1, e_2, \dots, e_m\}$. La matrice di incidenza \mathbf{F} registra un'entrata $f_{ij} = 1$ se il nodo del primo set a_i è legato al nodo del secondo set e_j e 0 altrimenti.

Un network *two-mode* di affiliazione, $\mathcal{N}(\mathcal{V}_1, \mathcal{V}_2, R)$, è composto da due set di nodi, \mathcal{V}_1 e \mathcal{V}_2 , appartenenti a classi di entità di natura diversa, connessi tra loro attraverso una relazione di affiliazione, o di partecipazione, R , che associa gli elementi del primo set a quelli del secondo set. Tale network sociale, come si è visto, può essere descritto formalmente attraverso il concetto più ampio di ipergrafo che consente di esprimere la ripartizione del set di nodi indotta dai (set) di legami di affiliazione, sottolineando come il singolo legame possa coinvolgere

subset di nodi di qualsiasi dimensione. I nodi inclusi in ogni subset sono connessi indirettamente attraverso la condivisione della relazione di appartenenza, co-appartenenza(101). Il concetto di ipergrafo, e di duale dell'ipergrafo, riflette l'idea, alla base dei network di affiliazione, che i soggetti possano essere inglobati in strutture più ampie (i subset di legami possono essere dati dall'appartenenza a gruppi, organizzazioni, comunità). Allo stesso tempo, le strutture sociali sono viste come emergenti dall'azione degli individui che ne fanno parte. Questa idea di poter analizzare le unità attraverso le entità a cui appartengono e, viceversa, le entità a partire dalle unità che ne fanno parte è stata anche alla base degli sviluppi teorici sul concetto sociologico di dualità delle persone e dei gruppi di Breiger(43). Un ipergrafo, $\mathcal{H}(\mathcal{V}, \mathcal{E})$, con un set finito di nodi $\mathcal{V} = \{v_1, v_2, \dots, v_n\}$ e un insieme di set legami (hyperedge) $\mathcal{E} = \{E_1, E_2, \dots, E_m\}$ con $E_j \in \mathcal{V}$ per $j = 1, \dots, m$, può essere rappresentato attraverso una matrice di incidenza, \mathbf{F} , di dimensioni $n \times m$, dove le righe rappresentano gli elementi del primo set di unità $\mathcal{V}_1 = \{a_1, a_2, \dots, a_n\}$ e le colonne rappresentano i singoli subset del secondo set di entità $\mathcal{V}_2 = \{e_1, e_2, \dots, e_m\}$. La matrice \mathbf{F} ha come entrata $f_{ij} = 1$ se il nodo a_i è incluso nel set di legami E_j e 0 altrimenti. La matrice di incidenza associata al *dual hypergraph* $\mathcal{H}^*(\mathcal{E}, \mathcal{V})$ è data dalla trasposta della matrice di incidenza associata all'ipergrafo \mathcal{H} e consente di guardare la ripartizione dei legami indotta dai nodi che ne fanno parte (mostra quali punti (che rappresentano gli eventi) sono incidenti con quali legami (che rappresentano attori) (21).

Dalla matrice di incidenza, o matrice *two-mode*, \mathbf{F} , è possibile ricavare due matrici di adiacenza, matrici *one-mode*, rispettivamente caso-per-caso derivante da $\mathbf{F}\mathbf{F}'$ e affiliazione-per affiliazione derivante da $\mathbf{F}'\mathbf{F}$, entrambe sono matrici non dirette e pesate. L'analisi di queste matrici consente di esplorare da un alto, i legami esistenti tra gli attori sulla base della loro comune appartenenza a collettività e, dall'altro, i legami esistenti tra le collettività sulla base della loro comune condivisione di alcuni attori.

La struttura di dati che deriva dalla matrice di incidenza, nodi-per-affiliazioni, è assimilabile a quella di una classica matrice casi-per-variabili, e ad essa si farà ampiamente ricorso nel seguito della trattazione indicandola come matrice *two-mode* o matrice di affiliazione.

Le matrici possono assumere forme complesse in relazione alla natura e al modo in cui le informazioni sui legami sono state misurate e codificate (192) (176) (124). Principalmente è possibile che i legami siano connotati da una o più delle seguenti caratteristiche: (i) direzione (non diretto/diretto); (ii) valenza (segno (+ -)); (iii) forza della relazione, intensità, molteplicità (valori).

1.4 Misure di distanza per dati relazionali

L'organizzazione dei dati relazionali in matrice rende possibile analizzare le caratteristiche strutturali del network, degli individui e gruppi che lo compongono, attraverso sia i modelli della teoria dei grafi sia modelli dell'algebra matriciale. Il concetto di distanza è alla base di numerosi sviluppi nella SNA orientati ad analizzare le proprietà strutturali dei punti/nodi (indici di centralità locali e globali) dei gruppi (*clique*, *n-clique*, *k-plex*, ecc.) e della struttura nel suo complesso (*Multidimensional Scaling*, *Clustering*). La nozione di distanza è fondamentale nel discorso della teoria relazionale nella misura in cui attraverso le sue declinazioni è possibile esprimere le forme di integrazione e differenziazione sociale in termini di somiglianza/differenza tra le unità interconnesse del sistema sociale. Il passaggio da un utilizzo euristico ad uno analitico del concetto di distanza ne richiede una definizione appropriata che implica la specificazione di un criterio globale che misuri la similarità/dissimilarità tra le unità. È possibile definire una funzione d che associa un numero reale positivo o nullo a ciascuna coppia di unità (i, j) appartenenti allo spazio considerato sulla base di determinate condizioni. Per tutte le unità v_i , v_j e v_u :

- (i) $d(i, j) \geq 0$, (non-negatività)
- (ii) $d(i, j) = 0$, se, e solo se, $i = j$; (separabilità)
- (iii) $d(i, j) = d(j, i)$; (simmetria)
- (iv) $d(u, j) \leq d(u, i) + d(i, j)$ (disuguaglianza triangolare)

Si parla di dissimilarità se si verificano la prima, la seconda e la terza condizione; si parla di metrica o distanza se valgono tutte le quattro condizioni. L'applicabilità

di una delle forme di misura è strettamente legata alla natura dei dati considerati (variabili numeriche, frequenze, variabili categoriali). Inoltre, la selezione di una misura rispetto all'altra implica una differenza nelle operazioni eseguibili. Infatti, un indice di dissimilarità consente solo di eseguire delle operazioni di confronto tra le caratteristiche delle coppie, mentre una metrica, o distanza, grazie alla condizione di disuguaglianza triangolare, consente anche la definizione di una relazione d'ordine tra le distanze dei punti in esame (136).

Proprietà delle misure di distanza nel grafo. Come si è visto in precedenza, negli sviluppi teorici radicati nella teoria dei grafi, il concetto di distanza è legato alla connettività del grafo ed, in particolare, al concetto di geodetica, la lunghezza minima del sentiero (*path*) tra due nodi. In un grafo connesso, la distanza tra due punti v_i e v_j è data dal sentiero di lunghezza minima da v_i e v_j , $d(i, j)$. La distanza geodetica soddisfa tutte e quattro le proprietà di una metrica. Il diametro (*diameter*) di un grafo \mathcal{G} è la distanza massima tra due punti di \mathcal{G} . Se non esiste un percorso tra due punti (ossia essi appartengono a diverse componenti del grafo) la distanza geodetica è infinita (o non definita). Di conseguenza, se e solo se \mathcal{G} è un grafo non diretto e connesso, l'insieme dei punti \mathcal{V} di \mathcal{G} , dotati di una distanza geodetica $d(i, j)$, rappresenta uno spazio metrico. A questo proposito è da notare come il concetto di distanza in un grafo si differenzi dal concetto di distanza geometrica nello spazio euclideo. La distanza tra due punti nel grafo è basata sulla geodetica che essendo il sentiero più breve implica una concezione della distanza come prossimità fisica nel sistema topologico definito dal grafo stesso (distanza di percorso), con l'impossibilità di calcolare la distanza tra coppie di punti che non fanno parte dello stesso componente, poiché non sono raggiungibili. La distanza geometrica euclidea, invece, valuta la distanza tra due punti basandosi sul sistema di riferimento cartesiano e sull'individuazione delle coordinate che li definiscono nello spazio (a due dimensioni), indipendentemente dalla vicinanza fisica, e quindi dal vincolo topologico, consentendo la definizione di una distanza anche per punti appartenenti a sottoinsiemi differenti.

La nozione di distanza misurata in termini di distanza geodetica può essere applicata anche in un grafo diretto \mathcal{G}_d . L'unica differenza, rispetto al caso non orientato, è legata alla necessità di considerare solo i legami orientati nella stessa

direzione. Ne deriva che, in un grafo diretto, la geodetica tra due punti v_i e v_j può essere diversa se si considera il percorso da v_i a v_j o il suo inverso dal v_j a v_i , quindi $d(i, j) \neq d(j, i)$. Questa evidenza implica che in un grafo diretto la geodetica non è una distanza, ma una quasi-metrica, in virtù del rilassamento dell'assioma della simmetria(72). Le distanze tra le coppie di nodi di un grafo possono essere rappresentate attraverso una matrice D in cui $d(i, j)$ costituisce la lunghezza del sentiero (*path*) più breve tra il nodo i e j . Accanto alle misure di distanza sviluppate a partire dalla teoria dei grafi è possibile applicare ai dati relazionali le misure di distanza della tradizione statistica classica.

Misure di distanza per dati numerici. La distanza Euclidea rappresenta il tipo di distanza più utilizzata quando le unità sono descritte da variabili numeriche. Tuttavia insieme ad essa se ne possono individuare alcune tra le più comuni. Per le unità v_i e v_j , descritte da m variabili numeriche, si ha:

$$d(v_i, v_j) = \sqrt{\sum_{i=1}^m (v_i - v_j)^2}$$

Per la distanza di Manhattan, anche nota come *city block*, si ha:

$$d(v_i, v_j) = \sum_{i=1}^m |v_i - v_j|$$

Entrambe le precedenti possono essere considerate come un caso particolare della distanza di Minkowsky:

$$d(v_i, v_j) = \left(\sum_{i=1}^m |v_i - v_j|^r \right)^{\frac{1}{r}}, \quad r \geq 0$$

Per $r = 1$, si ha la distanza di Manhattan, e per $r = 2$ si ha la distanza Euclidea. La distanza di Minkowski possiede un'utile proprietà: maggiore è il valore di r , maggiore è l'influenza delle differenze $|v_i - v_j|$ più grandi sulla distanza tra le unità. Nel caso in cui $r = \infty$, la distanza Minkowski viene chiamata distanza di Chebyshev

$$d(v_i, v_j) = \max_i |v_i - v_j|$$

Un altro indice che può essere utilizzato come misura di similarità è il coefficiente di correlazione di Pearson:

$$r(v_i, v_j) = \frac{\sum_{i=1}^m (v_i - \mu_{v_i})(v_j - \mu_{v_j})}{\sqrt{\sum_{i=1}^m (v_i - \mu_{v_i})^2 \sum_{i=1}^m (v_j - \mu_{v_j})^2}}$$

Con $\mu_{v_i} = \frac{1}{m} \sum_{i=1}^m v_i$ e $\mu_{v_j} = \frac{1}{m} \sum_{i=1}^m v_j$

Misure di distanza per dati di frequenza. Data una tabella di contingenza in cui sono riportate le frequenze associate a ciascuna modalità delle variabili, la i -esima riga, o j -esima colonna, costituisce il profilo dell'insieme di unità che possiede la modalità i della prima variabile, e j della seconda. Il tal caso la metrica più appropriata è la distanza del chi-quadrato. La distanza tra due punti riga v_i e $v_{i'}$ è data da:

$$d^2(v_i, v_{i'}) = \sum_j \frac{1}{f_{.j}} \left(\frac{f_{ij} f_{i'j}}{f_{i.} f_{i'.}} \right)^2$$

Mentre la distanza tra due punti colonna v_j e $v_{j'}$ è data da:

$$d^2(v_j, v_{j'}) = \sum_i \frac{1}{f_{i.}} \left(\frac{f_{ij} f_{ij'}}{f_{.j} f_{.j'}} \right)^2$$

Indicando con f_{ij} le frequenze relative della tabella, $f_{i.}$, i marginali di riga ed $f_{.j}$ i marginali di colonna. La distanza del chi-quadrato può essere considerata come una distanza Euclidea ponderata, con fattori di ponderazione uguali ai reciproci dei marginali di colonna, nel caso della distanza tra due punti riga, o a quelli di riga, nel caso di distanza tra due punti colonna (136) .

Misure di distanza per dati binari. Molte misure di similarità sono state definite per le unità descritte da variabili binarie. Esse sono determinate principalmente dalle frequenze della tabella di contingenza per una coppia di unità per la quale si misura la somiglianza. La tabella di contingenza per le unità v_i e v_j dove i valori di tutte le m variabili sono 1 e 0 è data da:

All'interno dello schema, la somma di tutte e quattro le frequenze corrisponde al numero delle variabili ($a + b + c + d = m$). La frequenza a conta la presenza

	1	0	
1	a	b	a+b
0	c	d	c+d
	a+c	b+d	m

Tabella 1.1: Schema per il calcolo degli indici di Similarità

congiunta di risposte positive, 1, per le unità v_i e v_j , e d conta la presenza congiunta di risposte negative, 0. Le frequenze b e c contano il numero di variabili per le quali le unità hanno fornito risposte diverse (9)(83). Diverse sono le misure sviluppate in letteratura, si riportano alcune tra le più utilizzate.

L'indice di similarità di Sokal–Michener (187) :

$$S_{SM}(v_i, v_j) = \frac{a + d}{a + b + c + d}$$

L'indice di similarità di Sokal–Sneath (162):

$$S_{SS}(v_i, v_j) = \frac{2(a + d)}{2(a + d) + b + c}$$

L'indice di similarità di Russell–Rao (174):

$$S_{RR}(v_i, v_j) = \frac{a + d}{a + d + 2(b + c)}$$

L'indice di similarità di Jaccard (130):

$$S_J(v_i, v_j) = \frac{a}{a + b + c}$$

Capitolo 2

Metodi di Clustering per dati relazionali: il Blockmodeling

2.1 Gruppi coesi e classi equivalenti

Una linea di sviluppo molto importante nella SNA è dedicata all'individuazione nel network sociale di sotto-strutture (subset) come entità rilevanti dotati di particolare significato. Tutti gli approcci adottati tentano di fornire una precisa definizione in termini strutturali del concetto di gruppo sociale, uno dei concetti che è stato oggetto di studio sin dagli albori della SNA, a partire dalle proprietà dei subset di attori all'interno del network (106, 192). L'approccio più immediato è quello che esamina le relazioni esistenti e distingue i gruppi in base al loro grado di coesione, secondo il livello di densità dei legami tra i membri del sottoinsieme. Tuttavia, gli individui che condividono le stesse relazioni possono essere raggruppati anche in assenza di legami condivisi. Questa differenza nel criterio alla base dell'aggregazione degli attori in sottogruppi rappresenta una delle distinzioni concettuali tra l'approccio relazionale e quello posizionale (56). L'approccio relazionale propone modelli basati sull'intensità delle relazioni e sul principio di coesione, come quello di gruppo primario e di *clique*. L'approccio posizionale sviluppa modelli basati sulla struttura delle connessioni tra gli attori e sul principio di equivalenza strutturale, come i modelli di status e *role-set* concepiti come posizioni occupate congiuntamente nel network. Pertanto, la ricerca di sotto-strutture rilevanti può seguire due criteri in cui l'obiettivo è quello

di individuare i modelli strutturali di connessioni che legano gli attori sociali in termini di sottogruppi coesi (insieme di attori fortemente legati tra loro) oppure di posizioni sociali (insieme di attori collegati in modo simile nel sistema sociale totale, non necessariamente connessi tra loro) (56, 102, 105, 106). Nel primo caso, le misure in uso sono state sviluppate, principalmente, sulla base delle proprietà del network (individui e gruppi al suo interno) a partire dai concetti della teoria dei grafi. Nel secondo caso, esse sono state sviluppate a partire da un approccio basato sull'algebra matriciale orientato a cogliere le proprietà strutturali delle posizioni sociali (status) e i modelli comportamentali che sono associati ad esse (ruoli)(69, 177, 178). Di seguito si discuteranno alcune delle idee centrali nell'analisi dei sottogruppi coesi, prima di affrontare in modo più dettagliato l'analisi delle sottostrutture come classi di attori equivalenti¹.

2.1.1 Metodi di Clustering per gruppi coesi

Nell'individuazione delle strutture di gruppo, sottostrutture rilevanti nel network sociale, è possibile seguire due direzioni guardando il network dal basso verso l'alto, approccio *bottom-up*, e dall'alto verso il basso, approccio *top-down*(123, 124). Il primo approccio parte dalla diade come struttura elementare e concepisce la struttura come emergente dalla sovrapposizione di entità più piccole, mentre il secondo parte dal network nel suo insieme e, al suo interno individua sotto-strutture, frammentandolo in unità più piccole (123, 124). Partendo da una prospettiva *bottom-up*, molti degli approcci orientati alla comprensione della struttura di gruppo di un network si basano sull'assunto in base al quale l'emergere del consenso sociale e di caratteristiche comuni tra gli attori sociali sia favorita dalla densità delle relazioni in cui essi sono coinvolti. Uno dei concetti più importanti sviluppati in letteratura è quello del sottogruppo coeso(56, 107). I sottogruppi coesi sono sottoinsiemi di attori tra i quali esistono legami relativamente intensi e frequenti. La rilevanza teorica di questo concetto risiede nel ruolo attribuito

¹In questo contesto ci si limiterà all'illustrazione dei concetti principali, ristretta al caso di network *one-mode* non orientate, per una trattazione più estensiva che tenga conto dei vari approcci sviluppati in relazione alle varie strutture di dati si rimanda a Wasserman e Faust (192), Newman e Girvan (161), Hanneman e Riddle (123, 124), Fortunato (100), Freeman (106).

alle forze sociali nel veicolare la coesione sociale e il consenso attraverso il contatto diretto e indiretto tra i membri del sottogruppo e attraverso una densità interna delle interazioni maggiore rispetto all'esterno. Molte sono le proprietà di un network sociale legate all'esistenza di sottogruppi coesi e diverse sono anche le possibili definizioni dei sottogruppi stessi a seconda della struttura del network da analizzare. Infatti, nel caso dei *one-mode* i sottogruppi coesi sono definiti a partire dalle relazioni dirette tra coppie di attori e dalle proprietà dei loro legami, mentre nei network *two-mode* i sottogruppi coesi sono definiti a partire dai legami esistenti tra gli attori indotti dall'appartenenza o dalla partecipazione congiunta ad organizzazioni o ad attività (192). Il concetto di sottogruppo coeso è stato operativizzato attraverso strumenti teorici radicati nella teoria dei grafi alla base dei quali c'è l'idea di distanza minima (geodetica) e di raggiungibilità. Uno dei primi concetti, e anche il più noto, è quello di *clique*. Una *clique* è un insieme di attori in un network legati tra loro da relazioni forti. In caso di relazioni non dirette, o bidirezionali, in termini matematici una *clique* è definita come massimo sottografo completo composto da tre o più nodi. Essa è costituita da un sottoinsieme di nodi, tutti adiacenti tra loro, e non esistono altri nodi adiacenti a tutti i suoi membri. Così definito il concetto di *clique* è molto restrittivo, in quanto implica l'impossibilità di appartenere a più subset contemporaneamente. Tuttavia, nei network di grande dimensione è possibile individuare pochi gruppi che corrispondono ad una *clique* in senso stretto ma diverse *clique* che in parte si sovrappongono. Molti sono i concetti che sono stati sviluppati per superare le restrizioni del concetto originario: *n-clique*, *n-clan*, e *k-plex*. Un *n-clique* è un massimo sottografo in cui la distanza geodetica maggiore tra due nodi non è maggiore di n , dove n è la lunghezza del percorso consentito a tutti i membri per creare una connessione. In questo primo caso si rilassa il concetto di *clique* ammettendo la possibilità di distanze maggiori di 1. In una *2-clique* ogni membro può raggiungere tutti gli altri attraverso massimo un intermediario. Nel concetto di *n-clan* vengono identificate *n-clique* nel network e vengono eliminate quelle che hanno un diametro maggiore di n . In questo modo si limita la possibilità di connessione consentendola ad una distanza geodetica non maggiore di n e solo all'interno del sottografo, per evitare che un membro dell'*n-clique* sia al di fuori della *clique* stessa. L'approccio del *k-plex* è un altro modo di rilassare il concetto

di *clique*; esso è basato sul *degree* del nodo e richiede che tutti i membri del sottogruppo siano adiacenti ad un numero minimo di altri membri del sottogruppo. Un *k-plex* è un massimo sottografo contenente n nodi, in cui ogni nodo è adiacente a non meno di $n - k$ nodi nel sottografo. L'analisi *k-plex* tende a trovare gruppi sovrapposti. Un ultimo approccio è quello del *k-core*, un gruppo massimo di attori, i quali sono collegati a qualche numero k di altri membri del gruppo. In un grafo non orientato, un *k-core* è un massimo sottografo connesso in cui il *degree* di tutti i nodi $> k$. Tale sottografo deve contenere almeno $k + 1$ nodi. Pertanto, un *k-core* è una parte di un network in cui ogni membro è legato ad almeno altri k . Quindi tale gruppo ha un certo grado di coesione senza avere necessariamente le caratteristiche di una *clique* (69, 123, 124, 192). Gli approcci sin qui trattati si basano sulle proprietà dei legami all'interno del sottogruppo (adiacenza, distanza geodetica, numero di legami tra i membri dei sottogruppi). Tuttavia, i sottogruppi coesi si possono caratterizzare anche per il modo in cui si differenziano dall'esterno. Un sottogruppo coeso, infatti, tende ad essere composto da soggetti che sono connessi in misura maggiore tra di loro che con membri esterni (179)(192). È possibile individuare altri aspetti legati al concetto di un sottogruppo coeso come la concentrazione dei legami all'interno del sottogruppo e il confronto tra l'intensità dei legami all'interno del sottogruppo quella dei legami al suo esterno. Inoltre, diviene rilevante considerare le proprietà dei legami tra membri all'interno del sottogruppo in confronto alle proprietà dei legami degli attori esterni ad esso. L'analisi di queste caratteristiche, in una prospettiva *top-down*, porta a spostare l'interesse su aree del network che presentano specifiche peculiarità, come la presenza di punti deboli nella struttura complessiva. Uno dei primi concetti proposti è quello di componente (*component*). Un componente completo rappresenta un sottogruppo a massima coesione in cui tutti i membri sono connessi tra loro all'interno e non hanno legami con i gruppi esterni. Ciò significa che tutti i membri sono adiacenti tra loro. Anche per questo concetto, come per quello di *clique*, sono stati formulati approcci che ne rilassano le restrizioni. In un grafo con legami orientati è possibile distinguere tra componenti deboli (*weak component*) e componenti forti (*strong component*). Nel primo caso si richiede che gli attori siano connessi, nodi adiacenti, senza la necessità di specificare la direzione del legame. Nel secondo caso, per essere considerati parti

del componente, è necessario che ci sia un sentiero diretto tra i due attori. Per analizzare la fragilità della struttura del network può essere utile l'approccio del *bi-component*. Tale metodo individua la presenza di punti di rottura (*cut-point*), ossia dei nodi che, se rimossi, portano alla disconnessione del grafo. Un attore che ricopre tale posizione potrebbe avere un ruolo particolare all'interno del network. Metodi alternativi di individuare sottostrutture nel network a partire da aree di vulnerabilità sono l'approccio del *lambda-set* e dei *bridges*. In questi approcci l'attenzione è posta sui legami, piuttosto che sui nodi, la cui rimozione genera la disconnessione del grafo. In questo caso, quindi, viene considerata la *line-connettivity*, la misura in cui una coppia di nodi rimane legata da qualche percorso, anche quando i legami vengono eliminati dal grafo, ed è basata sulla *line-cut*, il numero minimo di linee che deve essere rimossa dal grafo per eliminare ogni percorso tra i due nodi. Un insieme di nodi è un *lambda-set* se ogni coppia di nodi al suo interno ha una *line-connettivity* maggiore di qualsiasi altra coppia di nodi costituita da un nodo all'interno del *lambda-set* ed un altro al suo esterno. Un altro approccio orientato all'identificazione di sottostrutture è basato sul concetto di modularità (160). Il criterio è quello di individuare dei gruppi che siano maggiormente legati all'interno piuttosto che all'esterno, nello specifico è dato dal rapporto tra i legami interni al gruppo e quelli con l'esterno comparato con il valore atteso di tale rapporto in caso di distribuzione casuale dei legami. Questo approccio ha il vantaggio di cercare di ridurre il numero di *line-cut* tra gruppi (123, 124, 192). Un approccio più esplorativo all'analisi della coesione nel network che cerca di rappresentare la struttura di gruppo dell'intero network si basa sull'algebra matriciale. I legami di un insieme di attori vengono analizzati attraverso una riorganizzazione della matrice di adiacenza in cui sono stati organizzati i dati, attraverso una operazione di permutazione delle righe e delle colonne. Attraverso questa operazione, gli attori che sono legati tra loro saranno vicini nella matrice e troveranno collocazione lungo righe e colonne vicino alla diagonale principale, mentre quelli che non sono legati tra loro saranno collocati nel triangolo in basso a sinistra o in quello in alto a destra della matrice di adiacenza. Tale approccio può essere applicato anche ad altri tipi di matrici e costituisce la base delle tecniche di Blockmodeling che verranno ampiamente affrontate nella prossima sezione. Anche i metodi dell'analisi statistica tradizionale sono state

impiegate per analizzare la struttura del network. Approcci in questa tradizione usano lo Scaling Multidimensionale, la Classificazione Gerarchica, l'Analisi Fattoriale principalmente per rappresentare e visualizzare la prossimità tra gli attori nel network(105, 106, 111, 112, 113, 192). Lo Scaling Multidimensionale(64) cerca di rappresentare la prossimità (similarità o dissimilarità) tra un insieme di entità in uno spazio di dimensioni ridotte in modo da riprodurre la struttura di prossimità della matrice originaria, ossia che le entità più prossime nei dati osservati siano più vicine nello spazio e viceversa. La struttura di dati iniziali più comune è costituita da una matrice simmetrica *one-mode* in cui sono riportati i valori di una misura di prossimità tra coppie di soggetti. Nell'analisi dei sottogruppi coesi una misura di distanza è data dalla distanza geodetica, che rappresenta la prossimità tra coppie di punti nel grafo. In questo modo le visualizzazioni prodotte attraverso lo Scaling Multidimensionale mostrano una rappresentazione dei dati relazionali in cui la vicinanza tra sottogruppi di attori può essere letta in termini di Teoria dei Grafi. Le tecniche di Scaling Multidimensionale sono state utilizzate anche in combinazione con quelle di classificazione utilizzando le misure di rete come struttura di base per analisi(176, 192). Anche le tecniche di Analisi Fattoriale come l'Analisi in Componenti Principali e l'Analisi delle Corrispondenze sono state ampiamente utilizzate nella SNA per esaminare la struttura di gruppo del network. Sono state utilizzate per analizzare direttamente matrici di adiacenza; indirettamente matrici di correlazione o covarianza derivate da matrici di adiacenza; direttamente una matrice di incidenza (network *two-mode*); indirettamente le due matrici di adiacenza (*co-membership*, o *co-affiliation*, e *event overlap matrix*) derivate da una matrice di incidenza (43, 94, 105, 171). È necessario sottolineare che nella maggior parte delle prime applicazioni queste tecniche di analisi multivariata sono state utilizzate, principalmente, come strumento di visualizzazione, trascurandone il significato statistico che ne è alla base (113). Sviluppi più recenti, invece, tendono a sottolinearne le potenzialità analitiche derivanti da un utilizzo incentrato su una corretta specificazione della struttura dei dati e della tecnica in funzione del problema conoscitivo (71, 112, 113, 169). Altri approcci recenti utilizzati per l'analisi della struttura di gruppo di un network sociale provengono dal filone dei fisici e dei *computer scientists* che definiscono

tale obiettivo conoscitivo come un problema di *Graph partitioning*, *Graph clustering*, o *Community detection* e hanno sia ridefinito alcuni approcci già esistenti sia sviluppati di nuovi con una particolare attenzione allo sviluppo di algoritmi e procedure per la gestione di network di grandi dimensioni, della complessità computazionale che ne deriva e dell'evoluzione temporale (100, 159, 160, 161).

2.2 Posizioni e Sistemi di Ruoli

Gli approcci per l'analisi della struttura di gruppo nel network illustrati sin ora sono accomunati dall'idea di aggregare i nodi in base a qualche definizione di prossimità (adiacenza, raggiungibilità, ecc.) (26, 151, 192, 197, 198). È possibile, tuttavia, che attori condividano caratteristiche strutturali simili pur non essendo direttamente in contatto tra loro. Quello che li accomuna non è tanto la vicinanza, ma soprattutto la condivisione di un modello di connessioni con altri attori nel network. Teoricamente, due attori occupano la stessa posizione se i loro modelli di interazione sono uguali. Attori strutturalmente simili vengono concepiti come capaci di svolgere la "stessa parte" nel network, ossia svolgono simili ruoli sociali. Questa idea trova ampio spazio nella teoria sociologica, antropologica e psicologica attraverso i concetti di struttura, posizione e ruolo. A partire da Simmel (184) una posizione sociale può essere concettualizzata come la forma oggettiva di un insieme di relazioni. Ad essa è strettamente legato il concetto di struttura che costituisce la base di quello di ruolo che può, a sua volta, essere definito come l'insieme delle norme e delle aspettative sociali associate ad un individuo che occupa una determinata posizione(62). I concetti di struttura e ruolo sociale, anche se talvolta non vengono definiti in modo univoco(192, 200), sono centrali anche nel lavoro di Linton (141), Nadel (157), Merton (150) che ne danno le prime definizioni. Di particolare importanza è il contributo di Nadel (157), punto di riferimento dei lavori fondanti di White, Boorman e Breiger(198) e di Boorman e White (23). Nel lavoro di Nadel la struttura sociale emerge dalla regolarità nei modelli di relazioni tra entità concrete. Essa non è concepita come un insieme di norme e valori astratti o come una classificazione di entità concrete in base ai loro attributi. Inoltre, nella descrizione della struttura sociale diviene necessario aggregare queste regolarità in modo coerente con la loro

natura reticolare. In quest'ottica "La società è un network di relazioni tra gli individui, e la struttura sociale è il network sottostante di relazioni tra ruoli o posizioni" (28, p. 17). L'approccio posizionale, nella SNA, parte dal presupposto che la struttura di ruolo di un gruppo e le posizioni degli individui nel gruppo possano essere analizzate a partire dalle legami misurati sui soggetti che ne fanno parte (56, 96). Nella SNA il concetto di posizione (classe equivalente) si riferisce ad un insieme di individui che sono inseriti allo stesso modo nel network di relazioni. Mentre il concetto di ruolo si riferisce ai modelli di relazioni intercorrenti tra gli attori o tra le posizioni. I sottogruppi coesi più o meno indipendenti, come insieme di attori densamente connessi al loro interno e poco connessi con l'esterno, in quest'ottica, possono essere considerati come un caso particolare di posizione (classe di attori equivalenti), caratterizzati da uno specifico modello di connessioni. L'aggregazione degli attori, considerando i modelli di connessioni in relazione all'intero network, tende a sottolineare l'interdipendenza tra i gruppi presenti nel network (28, 56, 170). Nell'analisi della struttura di gruppo attraverso l'approccio posizionale e di ruolo è possibile seguire, congiuntamente o meno, due direzioni. Da un lato, l'identificazione delle posizioni sociali come insieme di attori simili nelle loro relazioni con gli altri (198), e, dall'altro, l'individuazione dei modelli di ruoli sociali come sistemi di relazioni tra attori e tra le posizioni (23). Un gran numero di lavori sono stati sviluppati seguendo l'una o l'altra direzione, in misura maggiore la prima (96, 192). Nella prima direzione si inserisce anche l'approccio sviluppato in questo lavoro, con una particolare attenzione agli ultimi sviluppi relativi all'analisi dei network *two-mode* attraverso l'analisi posizionale (77, 81, 83). Prima di procedere nella trattazione relativa al tema di interesse verranno illustrati i concetti base e le principali procedure dei metodi di Blockmodeling sviluppati specificamente per l'analisi di posizioni e ruoli. La forma più semplice di analisi posizionale parte da un network *one-mode* e da uno o più relazioni rilevate sugli attori che ne fanno parte. L'obiettivo principale dell'analisi posizionale standard (96, 192) è quello di individuare sottoinsiemi di attori che sono simili rispetto al loro modello di relazioni. La similarità, o corrispondenza strutturale (28), viene definita in termini di equivalenza tra gli attori rispetto a qualche proprietà matematica formale. Quest'ultima costituisce il criterio in base al quale gli attori vengono raggruppati insieme in una specifica

posizione. Attraverso tale procedura questi ultimi vengono assegnati a specifiche classi di equivalenza, ognuna delle quali è costituita da attori considerati "identici" (almeno teoricamente) rispetto a quella specifica proprietà. L'equivalenza strutturale rappresenta una delle proprietà matematiche formali utilizzate per la definizione delle classi di equivalenza. L'identificazione delle classi di equivalenza, o posizioni, e la descrizione delle relazioni esistenti tra essi viene condotta attraverso le tecniche di Blockmodeling.

2.3 Introduzione al Blockmodeling

Del Blockmodeling possono essere individuate diverse definizioni, ognuna delle quali permette di metterne in luce aspetti differenti. Come definizione di base, il Blockmodeling comprende un insieme di metodi che consentono, da un lato, la partizione delle unità di un network in classi equivalenti (posizioni o cluster) e, dall'altro, la partizione in blocchi dell'insieme dei legami intercorrenti tra tali classi equivalenti (54, 81, 142, 198). Il Blockmodeling cerca di individuare cluster di unità che hanno modelli di relazioni simili tra loro, inoltre, si pone l'obiettivo di interpretare il modello di relazioni intercorrenti tra cluster (10). In modo più generale, il Blockmodeling comprende un insieme di metodi per la partizione di un network sulla base di uno specifico criterio di equivalenza. Esso consente di perseguire uno dei principali obiettivi della SNA, ossia l'individuazione della "struttura fondamentale" di un network, consentendone l'analisi e facilitando la comprensione dei fenomeni di network (77, 82, 83). In aggiunta, il Blockmodeling può essere visto come un metodo di riduzione dei dati, capace di eliminare gli elementi ridondanti in un sistema osservato per produrne un modello semplificato di relazioni tra tipi unità (10, 16, 28). Da un punto di vista teorico e concettuale, infine, il Blockmodeling è visto come strumento per l'analisi delle posizioni occupate dagli attori in una struttura sociale e del sistema di ruoli ad esse associate (23, 28, 83, 98, 131, 198). Borgatti e Everett (28), tuttavia, sottolineano che quest'ultimo obiettivo non possa essere correttamente perseguibile nel caso si utilizzi come criterio di aggregazione quello dell'equivalenza strutturale. Esso, infatti, richiedendo che gli attori siano raggruppati insieme se hanno gli stessi tipi di legami con gli "stessi" attori, viene a configurarsi come un criterio di

equivalenza locale, piuttosto che globale che non cattura pienamente i concetti di posizione e ruolo. In quest'ottica il concetto più adatto a cogliere tale distinzione è quello di equivalenza regolare, che, come si chiarirà in seguito, richiede più genericamente che le unità, o attori, abbiano gli stessi tipi di legami con attori "equivalenti", ossia con lo stesso "tipo" di *alters* (28, p. 92).

2.3.1 Equivalenza Strutturale, Equivalenza Regolare, Equivalenza Generalizzata

Il Blockmodeling può essere definito come un insieme di procedure empiriche e si basa sull'idea che le unità in una network possono essere raggruppate in base alla misura in cui esse sono equivalenti in termini di qualche definizione significativa di equivalenza. L'utilizzo di criteri di equivalenza diversi porta a differenti partizioni del network (89) (83, p. 170). In letteratura è possibile individuare diverse definizioni di equivalenza, tuttavia, due approcci sono quelli più importanti, di cui il secondo è una generalizzazione del primo, e ruotano intorno a due idee centrali (82, 83):

- *Equivalenza Strutturale*. Unità equivalenti hanno lo stesso modello di connessioni con gli *stessi* vicini.
- *Equivalenza Regolare*. Unità equivalenti hanno lo stesso o simile modello di connessioni con *diversi* vicini.

Equivalenza Strutturale. La definizione di equivalenza strutturale è stata originariamente sviluppata da Lorrain e White (142). Per questi studiosi le unità strutturalmente equivalenti sono legate allo stesso modo con tutte le altre unità nel network. Formalmente, una permutazione $\varphi : \mathcal{V} \rightarrow \mathcal{V}$ costituisce un automorfismo della relazione R se e solo se (83, p. 171)(25):

$$\forall i, j \in \mathcal{V} : (iRj \Rightarrow \varphi(i)R\varphi(j))$$

Le unità i ed j sono strutturalmente equivalenti, $i \equiv j$, se, e solo se, la permutazione $\pi = (ij)$ è un automorfismo della relazione R . Pertanto, i e j sono strutturalmente equivalenti se e solo se:

- (i) $iRj \Leftrightarrow jRi$,
- (ii) $iRi \Leftrightarrow jRj$
- (iii) $\forall z \in \mathcal{V} \setminus \{i, j\} : (iRz \Leftrightarrow jRz)$
- (iv) $\forall z \in \mathcal{V} \setminus \{i, j\} : (zRi \Leftrightarrow zRj)$

Tale espressione in forma matriciale può essere espressa come: $x_i \equiv x_j$ se e solo se

- (i) $rij = rji$,
- (ii) $rii = rjj$
- (iii) $\forall k \neq \{i, j\} : r_{ik} = r_{jk}$
- (iv) $\forall k \neq \{i, j\} : r_{ki} = r_{kj}$

Nonostante la precisa formalizzazione, il concetto risulta fortemente restrittivo nelle applicazioni empiriche poiché impone la condizione che le unità debbano essere legate allo stesso modo alle *stesse* unità vicine (*"same" neighbor*). Di conseguenza, negli studi empirici è molto raro trovare casi di equivalenza strutturale "perfetta" (77, 83). La difficoltà di stabilire un blockmodel a partire da un criterio di equivalenza strutturale perfetta ha condotto alla ricerca di strategie per una sua approssimazione. Tali strategie si basano sulla richiesta che gli attori raggruppati insieme siano "abbastanza" strutturalmente equivalenti. Nel caso di un approccio di Blockmodeling indiretto, tale "rilassamento" assume varie forme a seconda della misura di (dis)similarità scelta per l'approssimazione. Tra le diverse definizioni operative sviluppate per approssimare l'equivalenza strutturale si ricordano quella basata sul coefficiente di Correlazione (45), e quella basata sulla distanza Euclidea (54). In tali formulazioni emerge una mancanza di strumenti formali che consentano di valutare quanto un blockmodel si adatti bene ai dati analizzati (83). Negli sviluppi dell'approccio diretto, l'introduzione di una misura di adeguatezza del blockmodel, rappresenta uno degli sforzi più significativi. A partire dalla definizione di equivalenza strutturale sono ammissibili quattro possibili blocchi diagonali ideali, strutture base per la realizzazione di un blockmodel

(7, 13, 82, 83): nullo, $b_{ij} = 0$; diagonale nullo, $b_{ij} = 1 - \delta_{ij}$; diagonale completo, $b_{ij} = \delta_{ij}$; completo, $b_{ij} = 1$. Dove δ_{ij} è la funzione delta di Kronecker con $\delta_{ij} = 0$ se $i \neq j$ e $\delta_{ij} = 1$ se $i = j$ e $i, j \in C$. I blocchi con solo 0 sono detti nulli mentre quelli con 1 sono detti completi.¹ Per i blocchi non diagonali, $R(C_i, C_j)$, $i \neq j$, sono permessi solo blocchi nulli e completi.

Equivalenza Regolare. La prima generalizzazione di equivalenza strutturale sviluppata è quella di equivalenza regolare (175, 195) in cui due unità sono regolarmente equivalenti se sono collegate in modo equivalente ad altre unità che sono a loro volta equivalenti. Sinteticamente, una relazione di equivalenza \approx su \mathcal{V} è un'equivalenza regolare se e solo se per ogni $i, j, z \in \mathcal{V}$, $i \approx j$ implica che (82, 83):

- (i) $iRz \rightarrow \exists w \in \mathcal{V} : (jRw \wedge w \approx z)$,
- (ii) $zRi \rightarrow \exists w \in \mathcal{V} : (wRj \wedge w \approx z)$

Data la natura e le proprietà di questo tipo di equivalenza, se $C = \{C_k\}$ è una partizione coerente con l'equivalenza regolare, $R(C_i, C_j)$, i tipi di blocchi corrispondenti, possono essere o nulli (vuoto) o composti da 1 (*one-covered*), o meglio con almeno un 1 in ciascuna delle righe e in ciascuna delle colonne (13, 83).

Equivalenza nei grafi k -partiti. Nel caso di grafi k -partiti, ossia nel caso in cui i legami nel grafo intercorrano solo tra classi di unità diverse e non siano ammessi legami tra unità della stessa classe (set disgiunti), vengono definite specifiche caratteristiche per le relazioni di equivalenza (83). Il grafo bipartito (network *two-mode*), che costituisce la struttura di riferimento per la presente proposta, è una delle possibili forme di un grafo k -partito in cui sono presenti due distinte classi di unità. La forma a blocchi possibile in tali grafi è quella in cui si hanno blocchi diagonali solo nulli e blocchi extra-diagonali, generalmente, non nulli. In un grafo k -partito completo, in cui tutte le coppie di vertici di cluster diversi sono legati da legami, la partizione determina una equivalenza strutturale esatta. Inoltre, il partizionamento di un grafo k -partito, dove ogni vertice $a_n \in C_i$ ha un legame con almeno un membro di ciascuno degli altri cluster C_j , determina una equivalenza regolare esatta.

¹Per ulteriori approfondimenti sui blocchi strutturali si rinvia a Doreian et al. (81)

Equivalenza Generalizzata. L'espressione di Equivalenza generalizzata viene utilizzata per indicare il tipo di relazione di equivalenza definita nell'ambito del Blockmodeling Generalizzato. Quest'ultimo si basa sulla definizione di tipi di blocchi ideali che corrispondono a specifici modelli di connessione (80). Quindi, tale concetto non definisce un tipo specifico di equivalenza, ma piuttosto una nozione guida per "la costruzione di equivalenze 'personalizzate'" (203, p. 20). L'Equivalenza generalizzata viene definita attraverso la specificazione dei tipi di blocchi (connessioni) permessi all'interno dei cluster e tra di essi ed, eventualmente, anche attraverso la specificazione delle posizioni. L'equivalenza generalizzata è definita direttamente dai tipi di blocchi consentiti nella definizione del blockmodel (80, 83). Come si è visto sino ad ora, l'equivalenza strutturale e regolare possono essere definite a partire da una serie di blocchi ideali consentiti e attraverso la definizione di blocchi ideali appropriati per queste equivalenze. Nello specifico, all'equivalenza strutturale corrisponde l'utilizzo di blocchi ideali completi e nulli, mentre all'equivalenza regolare corrisponda l'utilizzo di blocchi ideali regolari e nulli(7, 13, 80, 83). Questi sono solo alcuni dei tipi di blocchi ideali possibili, a cui corrispondono specifici modelli di connessioni, che possono essere ulteriormente estesi. Questa elasticità è tra i maggiori vantaggi del Blockmodeling Generalizzato, ossia la possibilità di definire nuovi tipi di blocchi ideali, radicati su modelli di connessioni concrete, a cui corrispondono nuovi tipi di equivalenze (83).

2.4 Differenti approcci al Blockmodeling

Tra l'insieme delle procedure di Blockmodeling è possibile distinguere due approcci, quello indiretto del Blockmodeling convenzionale, classico e quello diretto del quale fa parte il Blockmodeling Generalizzato. Negli approcci indiretti prima dell'esecuzione del blockmodel è necessario convertire i dati iniziali attraverso una misura di similarità o dissimilarità tra le unità del network compatibile con il criterio di equivalenza selezionato, e, in seguito, applica alla matrice di distanze una delle tecniche di clustering classiche allo scopo di individuare gruppi di unità equivalenti. Negli approcci diretti, invece, viene cercata, direttamente sui dati originari, la partizione che meglio si adatta al criterio di equivalenza selezionato. In questa procedura la qualità dell'adattamento della partizione ai dati viene

misurato attraverso una funzione criterio specifica (7, 13, 77, 80, 82, 83, 98). In letteratura possono essere individuati diversi strategie di approccio diretto alcune delle quali si sforzano di definire una funzione criterio specifica per un certo tipo di equivalenze, o per la ricerca di un certo tipo di proprietà (90, 103). Un ulteriore sviluppo in questa direzione è l'approccio del Blockmodeling Generalizzato (80, 82, 83) che, oltre alla specificazione di una funzione criterio integrata nel blockmodel, sviluppa una serie di concetti e strumenti in un approccio più ampio in cui la struttura empirica del network viene messa in relazione ad una struttura ideale definita a partire da molteplici modelli di connessioni ideali (tipi di blocchi ideali). Quest'ultimo costituirà una parte centrale del presente lavoro.

Tra le caratteristiche specifiche del Blockmodeling convenzionale si possono individuare (83): (i) utilizzo di un approccio indiretto; (ii) utilizzo esclusivo dell'equivalenza strutturale e regolare; (iii) mancanza di misure formali per la valutazione della qualità dell'adattamento delle partizioni.

Per contro, tra le caratteristiche principali del Blockmodeling Generalizzato si hanno: (i) Utilizzo di un approccio diretto che consente di evitare le possibili distorsioni introdotte dall'utilizzo di una misura di (dis)similarità; (ii) introduzione di un numero maggiore di blocchi, oltre a quelli nullo, completo e regolare; (iii) possibilità di adottare un approccio deduttivo. L'approccio indiretto può essere solo induttivo, e quindi esplorativo, in quanto, non essendo possibile specificare né il numero di cluster né la posizione dei tipi di blocchi, può solo cercare di individuare cluster e blocchi nella fase finale, dopo che i dati sono stati trasformati. Nell'approccio diretto, invece, è possibile la "pre-specificazione" (*pre-specification*), ossia un processo in cui la conoscenza, teorica e/o empirica, di cui è in possesso il ricercatore viene utilizzata per specificare alcuni aspetti quali la collocazione del tipo di blocco nel blockmodel. In un Blockmodeling Pre-specificato è possibile specificare la collocazione di alcuni o di tutti i tipi di blocchi (pre-specificazione completa), inoltre, è possibile inserire dei vincoli (*constraint*), parziali o totali, rispetto all'appartenenza delle unità a determinati cluster. La possibilità di poter inserire pre-specificazioni e vincoli consente di considerare il Blockmodeling anche come procedura deduttiva, quindi anche confermativa; (iv) costruzione del blockmodel a partire da una serie di tipi di blocchi ideali (o ammissibili) e valutazione dello "scostamento" del blockmodel empirico. La presenza

di 1 all'interno di un blocco ideale nullo, o di uno 0 in un blocco ideale completo, non sono ammissibili rispetto al tipo di blocco e costituiscono delle incongruenze. Queste ultime vengono definite "inconsistenze" e sono alla base del calcolo della funzione criterio, misura che consente di valutare direttamente la qualità della partizione. La funzione criterio sintetizza il numero delle inconsistenze e fornisce una misura dell'adattamento ai dati, ossia dello scostamento del blocco empirico dal tipo di blocco ideale. Di conseguenza, uno degli obiettivi del ricercatore, ma non la finalità assoluta, è quello di trovare delle partizioni, e quindi il blockmodel, in cui la funzione criterio assuma il minimo valore possibile. A tal proposito si ricorre a procedure di clustering di ottimizzazione locale (*local optimization clustering*) che cercano di individuare la migliore combinazione delle partizioni per cui la funzione criterio è minima. La funzione criterio assume valore 0 quando il blockmodel ideale modella perfettamente i dati empirici. È da sottolineare che ogni partizione individuata nel blockmodel stabilisce una relazione di equivalenza per le unità che ne fanno parte. Sintetizzando, l'approccio del Blockmodeling Generalizzato può essere definito come un processo in cui si ha la specificazione dei tipi di blocchi ideali, l'organizzazione degli stessi in un modello di blockmodel e, infine, l'applicazione del blockmodel ai dati relazionali (83). Tale approccio risulta estremamente flessibile in quanto consente di passare, con gradi diversi e anche nel corso dello stesso studio, da un approccio esplorativo, nel caso non si abbiano informazioni aggiuntive sull'oggetto di studio, ad uno confermativo, nel caso le informazioni aggiuntive consentano la specificazione, totale o parziale, dei vari aspetti legati alla definizione della struttura del Blockmodeling. Quest'ultimo punto risulta particolarmente significativo per il presente lavoro e costituirà uno degli sviluppi che verranno mostrati nel corso della trattazione.

2.4.1 Blockmodeling Convenzionale

Con l'espressione Blockmodeling Convenzionale viene indicato l'approccio indiretto tradizionale di Blockmodeling per distinguerlo dal Blockmodeling Generalizzato e, più in generale, dall'approccio diretto(7, 13, 80, 82, 83). Prima di definire formalmente e sinteticamente i tratti principali del Blockmodeling Convenzionale è necessario specificare alcune considerazioni di fondo. Come si è visto,

l'obiettivo del Blockmodeling, indipendentemente dal tipo di approccio, è quello di individuare, in un determinato network $\mathcal{N} = (\mathcal{V}, R)$, cluster (classi) di unità (attori) che sono uniti dalla condivisione di caratteristiche strutturali definite a partire da una relazione R . Ognuno dei cluster costituisce una posizione e le unità al loro interno hanno gli stessi o simili modelli di connessione. I Cluster definiscono una partizione $\mathbf{C} = \{C_1, C_2, \dots, C_k\}$, ossia un clustering delle unità \mathcal{V} , in cui ogni partizione \mathbf{C} determina una relazione di equivalenza e, viceversa, ogni relazione di equivalenza determina una partizione. Un clustering \mathbf{C} genera anche la partizione della relazione R in blocchi, $R(C_i, C_j) = R \cap C_i \times C_j$. Ogni blocco è definito dalle unità appartenenti ai cluster C_i e C_j ed è composto di tutti gli archi che vanno dalle unità nel cluster C_i alle unità nel cluster C_j . Se $i = j$, il blocco $R(C_i, C_i)$ è detto blocco diagonale. Un blocco, quindi, definisce "una relazione tra due cluster di unità" (83, p. 169). In un blockmodel con k posizioni (cluster) sono presenti K^2 blocchi, in cui k sono blocchi diagonali e $k(k - 1)$ sono blocchi extra-diagonali. "Un blockmodel è costituito dalle sottostrutture ottenute attraverso l'identificazione di tutte le unità dello stesso cluster, all'interno del clustering \mathbf{C} " (83, p. 169). Un blockmodel così definito può essere rappresentato anche attraverso un grafo ridotto oppure attraverso una matrice relazionale, detta "matrice immagine". All'interno del grafo ridotto i vertici rappresentano le posizioni, ciascuna delle quali è definita da un cluster di unità, e gli archi rappresentano le relazioni tra le posizioni definite dalla struttura dei rapporti tra i cluster di unità, ossia i blocchi. Una caratteristica essenziale è che il partizionamento è realizzato utilizzando esclusivamente sulla base delle informazioni strutturali contenute nella relazione R , e le unità nello stesso cluster sono equivalenti tra loro esclusivamente rispetto alla relazione R . In sintesi, il blockmodel è dato dalla matrice immagine, con le posizioni definite dai cluster e le relazioni tra loro intercorrenti. Pertanto, "stabilire blockmodel da dati relazionali empirici è il cosiddetto problema di Blockmodeling" (83, p. 170). Il problema di stabilire una partizione delle unità in un network (classi di unità), sulla base di un certo criterio di equivalenza, selezionato sulla base di conoscenze sostantive, teoriche e/o empiriche del ricercatore, viene considerato come un caso particolare del problema di clustering (10, 83). Nella sua forma generale un problema di clustering può essere espresso come segue (83). Dato un insieme di unità \mathcal{V} ,

l'obiettivo è quello di individuare sottoinsiemi di unità, cluster, \mathbf{C} , tali che siano omogenei, o ben distinti, sulla base delle variabili misurate. L'insieme dei cluster genera un clustering. La ricerca di tali cluster può essere formulata come un problema di ottimizzazione il cui l'obiettivo è determinare il clustering \mathbf{C}^* per il quale $P(\mathbf{C}^*) = \min_{\mathbf{C} \in \phi} P(\mathbf{C})$, dove \mathbf{C} è un clustering di un certo insieme di vertici, unità o attori \mathcal{V} , ϕ è il set di tutti i possibili clustering, e $P : \phi \rightarrow \mathbb{R}$ definisce una funzione criterio. Poiché il set di clustering possibili ha un numero finito, la soluzione del problema di clustering esiste sempre. Tuttavia, poiché questo insieme è generalmente di grandi dimensioni, non è facile trovare una soluzione ottimale. I clustering gerarchici e quelli basati sul partizionamento sono tra i più utilizzati per l'analisi dei dati relazionali. Un clustering $\mathbf{C} = \{C_1, C_2, \dots, C_k\}$ è una partizione dell'insieme di unità \mathcal{V} , se $\bigcup_i C_i = \mathcal{V}, \forall i \neq j \Rightarrow C_i \cap C_j = \emptyset$. Nel Blockmodeling Convenzionale la soluzione al "problema del Blockmodeling" viene ottenuta attraverso un approccio indiretto che prevede due fasi principali. In un primo momento è necessario individuare una misura di similarità o dissimilarità compatibile con il criterio di equivalenza selezionato e convertire i dati originari in una matrice di dissimilarità. In un secondo momento, tale matrice viene utilizzata come input per tecniche di analisi dati standard per la ricerca di cluster, sia clustering gerarchici sia non gerarchici. È possibile adottare come base un qualsiasi tipo di equivalenza, a condizione che esista una misura di (dis)similarità con essa compatibile. Tra le diverse procedure operative implementate che adottano un approccio indiretto le più rilevanti sono CONCOR (45) che prevede un'approssimazione del criterio di equivalenza strutturale a partire da una serie di correlazioni iterate, e STRUCTURE (54) basata sulla distanza Euclidea. Alla fine del processo si cerca di interpretare i cluster di unità ottenuti sulla base della relazione e degli attributi delle unità incluse nel cluster ¹. Come si è visto, in questo processo, è necessario definire una misura di (dis)similarità. Quest'ultima può essere definita a partire da proprietà strutturali delle singole unità (*degree*, numero di unità a distanza d dall'unità a , misure di centralità ecc.), o a partire dalle proprietà strutturali relative a coppie di unità (es. numero di vicini comuni, nodi adiacenti, delle unità a e b ; lunghezza del sentiero più breve

¹Per una trattazione più estesa e dettagliata sul sugli approcci indiretti si rimanda ai testi di (83, 96, 192)

tra a e b) (14, 83). Definita la misura di (dis)similarità è necessario verificarne la compatibilità con il tipo di equivalenza considerato. Successivamente è possibile procedere con le procedure di clustering. Una misura di dissimilarità d può essere considerata compatibile con una equivalenza \equiv , se e solo se per ogni coppia di unità $i, j \in \mathcal{V}$ $i \equiv j$, il che implica $d(i, j) = 0$ (83). Non tutte le misure di dissimilarità sono compatibili con l'equivalenza strutturale ¹

Misure di dissimilarità per l'equivalenza strutturale. Una proprietà $q : \mathcal{V} \times \mathcal{V} \rightarrow \mathbb{R}$ è strutturale se, per ogni automorfismo φ della relazione R , e per ogni coppia di unità $a, b \in \mathcal{V}$, vale : $q(a, b) = q(\varphi(a), \varphi(b))$. Sia $q(a, b)$ una proprietà strutturale delle unità a e b . Allora, se $x \equiv y$, si ha:

- (i) $q(x, y) = q(y, x)$,
- (ii) $q(x, x) = q(y, y)$
- (iii) $\forall k \in \mathcal{V} \setminus \{x, y\} : q(x, k) = q(y, k)$
- (iv) $\forall k \in \mathcal{V} \setminus \{x, y\} : q(k, x) = q(k, y)$

Se valgono le condizioni i – iv, la proprietà q è sensibile (sensitive), ciò implica che $x \equiv y$. Per una proprietà sensibile q , è possibile descrivere ogni unità a da un vettore $[a] = [q(a, x_1), q(a, x_2), \dots, q(a, x_n), q(x_1, a), \dots, q(x_n, a)]$ e definire una dissimilarità tra le unità $a, b \in \mathcal{V}$ come $d(a, b) = D([a], [b])$, dove D è una dissimilarità tra le descrizioni corrispondenti. Tra le misure di dissimilarità, e le loro versioni compatibili con il principio dell'equivalenza strutturale, si possono individuare (83, p. 181): Distanza di Manhattan:

$$d_m(x_i, x_j) = \sum_{s=1}^n (|q_{is} - q_{js}| + |q_{si} - q_{sj}|)$$

Distanza Euclidea:

$$d_E(x_i, x_j) = \sqrt{\sum_{s=1}^n ((q_{is} - q_{js})^2 + (q_{si} - q_{sj})^2)}$$

¹Per una trattazione più ampia e dettagliata Batagelj et al. (13); Doreian et al. (83, Cap. 6).

Distanza di Manhattan troncata:

$$d_s(x_i, x_j) = \sum_{s=1, s \neq i, j}^n (|q_{is} - q_{js}| + |q_{si} - q_{sj}|)$$

Distanza Euclidea troncata (92):

$$d_S(x_i, x_j) = \sqrt{\sum_{s=1, s \neq i, j}^n ((q_{is} - q_{js})^2 + (q_{si} - q_{sj})^2)}$$

Distanza di Manhattan corretta come dissimilarità $p > 0$

$$d_c(p)(x_i, x_j) = d_s(x_i, x_j) + p(|q_{ii} - q_{jj}| + |q_{ij} - q_{ji}|)$$

Distanza Euclidea corretta come dissimilarità (53):

$$d_e(p)(x_i, x_j) = \sqrt{d_S(x_i, x_j)^2 + p((q_{ii} - q_{jj})^2 + (q_{ij} - q_{ji})^2)}$$

Misura di dissimilarità corretta:

$$d_C(p)(x_i, x_j) = \sqrt{d_c(p)(x_i, x_j)}$$

Dove p è un qualsiasi valore positivo, in genere 1 o 2, che conta il numero di volte che la coppia diagonale corrispondente viene conteggiata. Tra queste misure di dissimilarità solo alcune sono compatibili (*compatible*) con l'equivalenza strutturale, ossia rispettano il presupposto in base al quale $x_i \equiv x_j \Leftrightarrow d(x_i, x_j) = 0$. Se una proprietà strutturale q è sensibile, la distanza di Manhattan corretta d_c , la dissimilarità corretta d_C e la distanza Euclidea corretta d_e sono compatibili con l'equivalenza strutturale. Ciò significa che due unità sono strutturalmente equivalenti esattamente quando ognuna di queste dissimilarità assume valore pari a zero. In questo caso $x_i \equiv x_j$, quindi le proprietà che definiscono la sensibilità di una proprietà strutturale assumono la seguente forma:

- (i) $q_{ij} - q_{ji} = 0$,
- (ii) $q_{ii} - q_{jj} = 0$
- (iii) $\forall k \neq i, j : q_{ik} - q_{jk} = 0$

$$(iv) \quad \forall k \neq i, j : qki - qkj = 0$$

In tal caso tutti i termini a destra delle formule nelle definizioni delle misure assumono valore 0 pertanto $d_c(x_i, x_j) = d_C(x_i, x_j) = d_e(x_i, x_j) = 0$. Tali affermazioni dimostrano che $x_i \equiv x_j \Rightarrow d(x_i, x_j) = 0$ e che $d(x_i, x_j) = 0 \Rightarrow x_i \equiv x_j$.

2.4.2 Dall'approccio indiretto all'approccio diretto

Nell'approccio indiretto vengono individuati alcuni punti critici fondamentali che hanno spinto alla formulazione di una serie di soluzioni alternative confluite in quello noto come approccio diretto al Blockmodeling (83). Un primo problema ravvisabile nell'approccio indiretto è legato alla necessità di definire una misura di dissimilarità compatibile con il tipo di equivalenza selezionato, e se questo può essere possibile per l'equivalenza strutturale, nel caso di quella regolare diviene più complicato. A tale problema l'approccio diretto risponde con una strategia più generale che non richiede l'uso di tali misure e agisce direttamente sui dati relazionali. Un secondo problema riguarda il ruolo e la definizione della funzione criterio. Quest'ultima negli approcci indiretti appare non strettamente associata allo specifico concetto di equivalenza alla base del Blockmodeling. Per far fronte a questo problema l'approccio diretto introduce una funzione criterio strutturata in modo da riflettere esattamente tale concetto. Un terzo problema ha a che fare con la mancanza di esplicite misure della qualità dell'adattamento del blockmodel ai dati. A tal proposito nell'approccio diretto è introdotta una strategia di ottimizzazione della funzioni criterio, definita appropriatamente, in modo che la sua minimizzazione diventi una misura diretta di quanto un blockmodel si adatti bene ai dati relazionali. Come si è visto, i modelli di connessioni associati all'equivalenza strutturale e regolare possono essere espressi in termini di tipi di blocchi ideali, blocco nullo (0) e completo (1) per la strutturale e blocco nullo e regolare (almeno un 1 per ciascuna riga e colonna) per l'equivalenza regolare. La costruzione della funzione criterio si basa proprio sulla disponibilità di questi tipi di blocchi ideali (possibili per una data equivalenza). Infatti essa, concretamente, misura lo scarto tra un certo blocco empirico e il corrispondente tipo di blocco ideale. Tale misura viene costruita per ogni blocco e la loro combinazione, in

genere addizione, va a costituire il valore finale della funzione criterio, la quale esprime proprio la misura della differenza tra il blockmodel empirico e quello ideale corrispondente. L'obiettivo dell'approccio diretto al Blockmodeling, pertanto, è quello di ricercare le partizioni di un network (cluster di unità e blocchi come cluster di legami) che minimizzino la funzione criterio (83).

2.4.2.1 Funzione criterio e procedura di clustering

Considerato un network *one-mode* basato su una singola relazione $\mathcal{N} = (\mathcal{V}, R)$ l'approccio diretto al Blockmodeling viene formalizzato a partire dalla definizione del concetto di inconsistenza del blocco, di un'appropriata funzione criterio e di una procedura di clustering che ne consenta la minimizzazione (14, 81, 83). La logica seguita nella costruzione della funzione criterio è quella di confrontare direttamente il clustering empirico con uno ideale caratterizzato da un modello di connessioni, interne ai cluster e tra cluster (ossia i blocchi), perfettamente aderenti al criterio di equivalenza considerato. Tale scelta è guidata dalla volontà di individuare una misura che "rifletta direttamente l'equivalenza considerata" (82, p. 83). Sia Θ l'insieme di tutte le relazioni relative al tipo di equivalenza selezionato (l'equivalenza strutturale, regolare, ecc.) su \mathcal{N} . Ogni relazione di equivalenza \sim su \mathcal{V} determina una partizione \mathbf{C} di \mathcal{V} , e viceversa. Sia Φ l'insieme di tutte le possibili partizioni corrispondenti alle relazioni in Θ . La costruzione di una funzione criterio $P(\mathbf{C})$ con le proprietà *P1*. $P(\mathbf{C}) \geq 0$ (non-negativa) e *P2*. $P(\mathbf{C}) = 0 \Leftrightarrow \sim \in \Theta$ (sensibile al criterio di equivalenza), consente di esprimere il problema di stabilire una partizione di un network, sulla base di un certo tipo specifico di equivalenza, come problema di clustering. Nel caso di equivalenze esatte, (*P2*), il valore minimo assunto da $P(\mathbf{C})$ è 0. Dato un clustering $\mathbf{C} = C_1, C_2, \dots, C_k$, sia $\mathcal{B}(C_i, C_j)$ l'insieme di tutti i blocchi ideali corrispondenti al blocco $R(C_i, C_j)$. In questo caso l'inconsistenza (o discrepanza) totale del clustering \mathbf{C} , rispetto ad clustering ideale, può essere espresso come somma delle inconsistenze all'interno dei blocchi e attraverso tutti i blocchi:

$$P(\mathbf{C}) = \sum_{C_i, C_j \in \mathbf{C}} p(C_i, C_j).$$

L'inconsistenza del blocco è data da

$$p(C_i, C_j) = \min_{B \in \mathcal{B}(C_i, C_j)} \delta(R(C_i, C_j), B)$$

dove il termine $\delta(R(C_i, C_j), B)$ misura la deviazione (numero di inconsistenze (83) del blocco $R(C_i, C_j)$ dal blocco ideale B , con una funzione δ compatibile con il tipo di equivalenza selezionato. A partire da queste premesse vengono definite le funzione criterio sia per l'equivalenza strutturale che per la regolare:

Funzione criterio per l'equivalenza strutturale. Dati i blocchi ideali ammissibili per l'equivalenza strutturale, nullo e completo, l'espressione $\delta(R(C_i, C_j), B)$ per i blocchi extra-diagonali viene essere espressa come

$$\delta(R(C_i, C_j), B) = \sum_{x \in C_i, y \in C_j} |r_{xy} - b_{xy}|.$$

Dove, r_{xy} è il legame osservato e b_{xy} è il valore corrispondente in un blocco ideale. La funzione criterio di $P(\mathbf{C})$ così definita è sensibile all'equivalenza strutturale:

$$P(\mathbf{C}) = 0 \Leftrightarrow \mathbf{C} \text{ definisce l'equivalenza strutturale.}$$

In questa definizione, lo scostamento tra i blocchi si basa sul presupposto che le inconsistenze $0 \rightarrow 1$ e $1 \rightarrow 0$ siano equiprobabili e abbiano la stessa importanza. Tuttavia, è possibile ponderare i due tipi di inconsistenza attraverso l'introduzione di due parametri, α , per penalizzare maggiormente gli 0 in un blocco di 1 (completo) e β per penalizzare maggiormente gli 1 in un blocco di 0 (nullo). In ultimo, si rileva una relazione tra numero di cluster e valore della funzione criterio, in quanto il valore di quest'ultima diminuisce all'aumentare del numero di cluster specificati per l'equivalenza strutturale (83).

Funzione criterio per l'equivalenza regolare. Dati i blocchi ideali ammissibili per l'equivalenza regolare, nullo e regolare, la funzione criterio si basa sulla misura di quanto il blocco sia regolare,

$$\delta(R(C_i, C_j), B) = \begin{cases} \text{numero di righe o colonne 1-covered,} & B \text{ è il blocco nullo} \\ \text{numero di righe o colonne 0,} & B \text{ è un blocco regolare} \end{cases}$$

la funzione criterio $P(\mathbf{C})$, così definita, è sensibile all'equivalenza regolare:

$$P(\mathbf{C}) = 0 \Leftrightarrow \mathbf{C} \text{ definisce l'equivalenza regolare.}$$

Procedura di clustering. Nell'ambito dell'approccio diretto, definita la funzione criterio appropriata al tipo di equivalenza, per risolvere il problema del Blockmodeling si ricorre ad una procedura di clustering di ottimizzazione locale (*local optimization clustering*), basata su un algoritmo di trasferimento (*relocation algorithm*). L'algoritmo parte da una partizione e procede ad una ricollocazione delle unità sino a quando la ricollocazione non genera una diminuzione ulteriore della funzione criterio. L'individuazione del vicinato (*neighborhood*) nella procedura di ottimizzazione locale, che porta alla ricollocazione e alla configurazione finale, si basa su due possibili trasformazioni: spostando un'unità x_k da un cluster C_i ad un altro cluster C_j (*transition*); scambiando due unità x_i e x_j tra cluster differenti C_i e C_j (*transposition*). Per ottenere una buona soluzione finale il procedimento viene ripetuto più volte con diverse partizioni casuali iniziali \mathbf{C} . È necessario notare che trattandosi di una procedura di ottimizzazione locale, non vi è la certezza che vengano individuate tutte le partizioni con una funzione criterio minima. Inoltre, si ottengono ottimi risultati su network di piccole dimensioni, (10, 15 unità), di conseguenza con l'aumento del numero delle unità è possibile che alcune partizioni ottimali non vengano individuate rendendo difficile l'applicazione del metodo per network di grandi dimensioni (83). Su questo aspetto si articolano diversi sviluppi che cercano sia di individuare funzioni criterio basate su definizioni differenti in grado di ridurre anche le complessità computazionali (61, 170), sia procedure alternative di ottimizzazione che cercano di garantire non solo un'ottimizzazione locale ma anche globale, soprattutto nell'ambito dell'estensione del Blockmodeling ai network *two-mode* (49, 50).

2.4.3 Blockmodeling Generalizzato

Il Blockmodeling Generalizzato costituisce una ulteriore estensione dell'approccio diretto in cui, oltre ai tipi di blocchi (nullo, completo e regolare) definiti per l'equivalenza strutturale e regolare vengono definiti nuovi tipi di blocchi ideali che portano alla definizione dell'equivalenza generalizzata. Questo processo è legato alla constatazione in base alla quale ogni partizione determina una relazione di equivalenza e viceversa (83). Pertanto, piuttosto che definire nuovi tipi di equivalenza ed derivarne i blocchi ideali possibili, vengono definiti direttamente

nuovi tipi di blocchi ideali che riflettono specifici modelli di connessioni. Questi nuovi tipi di blocchi ideali definiscono delle partizioni che inducono nuove forme di equivalenza. Questo processo porta ad un cambiamento nel modo di concettualizzare l'equivalenza e di stabilire il blockmodel. Va notato come ponendo l'attenzione sui tipi di blocchi, venga meno la necessità di misurare l'equivalenza per coppie di unità (83). Quello che risulta rilevante non è più quest'ultima ma il *pattern* complessivo di connessioni del blocco. Pertanto l'interesse si sposta sulla definizione dei tipi di blocchi ideali, sulle procedure empiriche per identificarli (nell'ambito dell'approccio diretto attraverso metodi di ottimizzazione), fino ad arrivare alla definizione di strategie in grado di impiegare le conoscenze aggiuntive nel processo di definizione e costruzione di vari tipi di blockmodel (a partire dalla pre-specificazione e dall'introduzione dei vincoli). Nelle Figure 2.1 e 2.2 sono rappresentati alcuni dei blocchi ideali che possono essere utilizzati nella descrizione dei legami tra i cluster (ossia le posizioni) di un blockmodel (e le relazioni all'interno di essi). I blocchi ideali illustrati sono relativi al caso di un network *one-mode* con una singola relazione (10, p. 6)(83, p. 2012)(82, p. 86).

	C_j		C_j		C_j
C_i	1 1 1 1 1	C_i	0 1 0 0 0	C_i	0 0 1 0 0
	1 1 1 1 1		1 1 1 1 1		0 0 1 1 0
	1 1 1 1 1		0 0 0 0 0		1 1 1 0 0
	1 1 1 1 1		0 0 0 1 0		0 0 1 0 1
	complete		row-dominant		col-dominant
	C_j		C_j		C_j
C_i	0 1 0 0 0	C_i	0 1 0 0 0	C_i	0 1 0 1 0
	1 0 1 1 0		0 1 1 0 0		1 0 1 0 0
	0 0 1 0 1		1 0 1 0 0		1 1 0 1 1
	1 1 0 0 0		0 1 0 0 1		0 0 0 0 0
	regular		row-regular		col-regular
	C_j		C_j		C_j
C_i	0 0 0 0 0	C_i	0 0 0 1 0	C_i	1 0 0 0 0
	0 0 0 0 0		0 0 1 0 0		0 1 0 0 0
	0 0 0 0 0		1 0 0 0 0		0 0 1 0 0
	0 0 0 0 0		0 0 0 1 0		0 0 0 1 1
	null		row-functional		col-functional

Figura 2.1: Tipi di blocchi ideali (82, 83)

I tipi di connessioni tra i cluster illustrati nelle immagini si basano su una serie di predicati (*predicate*), ossia descrizioni che devono essere rispettate dai blocchi $R(C_i, C_j)$. La definizione di alcuni nuovi tipi di blocchi ideali si basa sul rilassamento di alcune condizioni relative ai modelli di connessioni. In questo modo accanto ai tipi di blocchi ideali nullo, completo e regolare possono esserne

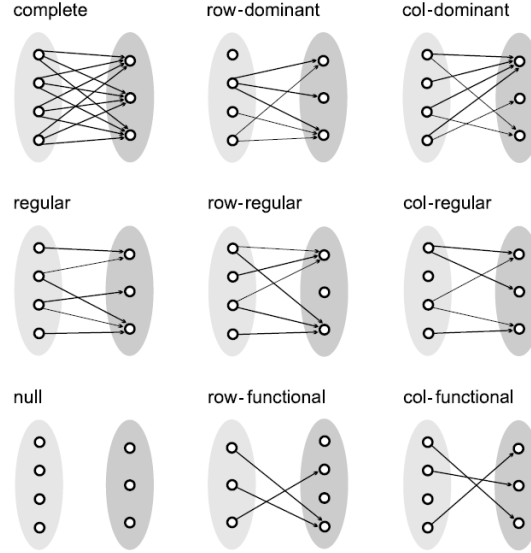


Figura 2.2: Tipi di connessioni tra blocchi (10) (83)

individuati altri, tutti basati su tipi di connessioni specifici. Sinteticamente, partendo dal primo blocco mostrato nella Figura 2.1, un blocco è *complete* se è tutto coperto da 1, coerente con l'equivalenza strutturale, se il blocco è diagonale ed ha la diagonale principale con tutti 0 tale blocco descrive una *clique*. Un blocco è *regular* se ha almeno un 1 in ciascuna riga e in ciascuna colonna, coerente con l'equivalenza regolare. Un blocco è *null* se è tutto coperto da 0, coerente con l'equivalenza strutturale e regolare. I tipi di blocchi ideali nuovi che seguono derivano da una generalizzazione legata all'indebolimento dell'equivalenza regolare (82, 83). Un blocco è *row-regular* se ha almeno un 1 in ciascuna riga, ed è *col-regular* se ha almeno un 1 in ciascuna colonna, generalizzazione legata all'indebolimento dell'equivalenza regolare. Un blocco è *row-dominant* se ha almeno una riga con tutte le celle che contengono 1 (almeno una unità del cluster C_i ha legami con tutte le unità del cluster C_j), ed è *col-dominant* se ha almeno una colonna con tutte le celle che contengono 1 (almeno una unità del cluster C_j è riceve legami da tutte le unità del cluster C_i). Un blocco è *row-functional* se ha esattamente un 1 in ogni riga (ogni unità del cluster C_i è legata ad una sola unità del cluster C_j), ed è *col-functional* se ha esattamente un 1 in ogni

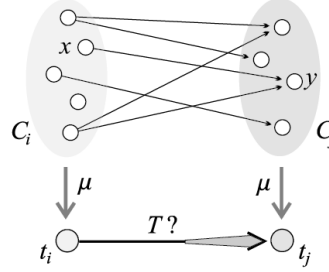


Figura 2.3: Blockmodeling Generalizzato (82, 83)

colonna (ogni unità del cluster C_j è legata ad una sola unità del cluster C_i)¹. A questi, possono essere aggiunti anche altri due tipi di blocchi ideali definiti sulla base della dimensione del blocco, il *degree density* γ e *degree bound* n . Il primo è basato su una soglia minima di densità γ ², il secondo su una soglia minima di unità n appartenenti al blocco. Definiti i tipi di blocchi ideali si può procedere alla formalizzazione del Blockmodeling Generalizzato come segue (10, 13, 82, 83) Dato un network, $\mathcal{N}(\mathcal{V}, R)$, con un set di unità \mathcal{V} e una relazione $R \subseteq \mathcal{V} \times \mathcal{V}$. Sia \mathcal{Z} un set di posizioni (o immagini) di gruppi di unità. Sia $\mu : \mathcal{V} \rightarrow \mathcal{Z}$ una mappatura che associa ogni unità alla sua posizione. Il cluster di unità $C(t)$ con la stessa posizione $t \in \mathcal{Z}$ è dato da $C(t) = \mu^{-1}(t) = \{x \in \mathcal{V} : \mu(x) = t\}$. Ne consegue che $\mathbf{C}(\mu) = \{C(t) : t \in \mathcal{Z}\}$ è una partizione ossia, un clustering, del set di unità \mathcal{V} . L'obiettivo generale è quello di individuare la natura del legame tra le due posizioni t_i e t_j (associate, rispettivamente, ai cluster C_i e C_j). È necessario individuare il tipo di legame che rispecchia la struttura del blocco $R(C_i, C_j)$. Un blockmodel (immagine o il grafo ridotto) è definito da $\mathbf{M} = (\mathcal{Z}, K, \mathcal{T}, \pi)$, dove \mathcal{Z} è un set di posizioni; $K \subseteq \mathcal{Z} \times \mathcal{Z}$ è un insieme di connessioni tra posizioni; \mathcal{T} è il set di predicati (predicate) utilizzati per descrivere i tipi di connessioni tra cluster in un network, a questo set appartiene anche il predicato *nul* che descrive il blocco nullo; una mappatura $\pi : K \rightarrow \mathcal{T} \setminus \{\text{nul}\}$ che assegna i predicati

¹Per una trattazione più dettagliata che tenga conto anche delle proprietà legate alle caratteristiche dei modelli di connessioni si rinvia a Doreian et al. (82, 83)

²Nell'approccio classico al Blockmodeling spesso viene fissata una soglia di densità minima per considerare il blocco come *oneblock*, corrispondente al blocco completo (Faust, Wasserman 1992)

alle connessioni, ossia il tipo di blocco ideale alle connessioni osservate. Una funzione o mappatura suriettiva $\mu : \mathcal{V} \rightarrow \mathcal{Z}$ determina un blockmodel \mathbf{M} del network \mathcal{N} se e solo se, soddisfa queste due condizioni: $\forall (t, w) \in K : \pi(t, w)(C(t), C(w))$ e $\forall (t, w) \in \mathcal{Z} \times \mathcal{Z} \setminus K : nul(C(t), C(w))$.

Il grafo ridotto può essere rappresentato anche attraverso una matrice \mathbf{M} detta matrice immagine o matrice modello del blockmodel in cui l'elemento generico è dato da m_{ij} = tipo di blocco per $R(C_i, C_j)$. Sia \sim una relazione di equivalenza su \mathcal{V} . Tale relazione partiziona il set di unità \mathcal{V} in cluster $[X] = \{y \in \mathcal{V} : x \sim y\}$. Si dice che la relazione \sim è compatibile con \mathcal{T} , o \mathcal{T} -equivalence, su un network \mathcal{N} se, e solo se, $\forall x, y \in \mathcal{V}, \exists T \in \mathcal{T} : T([x], [y])$. La nozione di compatibilità per $\mathcal{T} = \{nul, reg\}$ rappresenta la definizione di equivalenza regolare e quella per $\mathcal{T} = \{nul, com\}$ quella per l'equivalenza strutturale. Per un'equivalenza compatibile \sim la mappatura $\mu : x \mapsto [x]$ determina un blockmodel con $\mathcal{Z} = \mathcal{V} / \sim$.

Nell'ambito del Blockmodeling Generalizzato è necessario notare un cambiamento nell'interpretazione del concetto di equivalenza (83). I nuovi tipi di blocco, come in precedenza, determinano delle partizioni, ed i cluster ottenuti, quindi, costituiscono lo stesso delle classi di equivalenza. Le unità nello stesso cluster sono equivalenti in termini di \mathcal{T} -equivalence alla base della partizione. Nel caso dei blocchi *row-regular* e *col-regular* l'idea che le unità di un cluster siano equivalenti tra loro mantiene il suo significato. Tuttavia, nei blocchi *row-dominant* e *col-dominant* il termine equivalente deve essere usato con cautela. Infatti, per definizione, almeno una unità è diversa dalle altre, in termini di modello di connessioni, all'interno dello stesso cluster, quindi, non si può sostenere che siano equivalenti nel senso di strutturalmente identici. Le unità all'interno delle posizioni, quindi, sono sempre equivalenti in termini di relazione di equivalenza (e le sue classi di equivalenza) che ha determinato la partizione (83).

Definiti i tipi di blocchi ideali e il Blockmodeling Generalizzato la procedura individuata per la realizzazione del blockmodel è la stessa indicata per l'esecuzione dell'approccio diretto(82, 83). Anche in questo caso il problema del Blockmodeling viene definito come un problema di ottimizzazione in cui dato un set tipi di connessione, ossia di blocchi ideali $T \in \mathcal{T}$, e un blocco $R(C_i, C_j)$, si cerca di individuare il tipo di blocco T che maggiormente riflette il modello di connessioni

$R(C_i, C_j)$, ossia $\pi(\mu(C_i), \mu(C_j)) = T$. Pertanto, il problema di ottimizzazione viene formalizzato come:

$$\mathcal{B}(C_i, C_j, T) = \{B \subseteq K(C_i, C_j) : T(B)\} \quad (2.1)$$

Inoltre, viene definita come $\delta(C_i, C_j, T)$ lo scostamento del blocco $R(C_i, C_j)$ dal blocco ideale più vicino. Sulla base di questa definizione di deviazione, e delle caratterizzazioni che, specificate per ogni di tipo di blocco, permettono di calcolarla, viene definita l'inconsistenza del blocco e la funzione criterio da minimizzare. L'inconsistenza del blocco $\varepsilon(C_i, C_j)$ di $R(C_i, C_j)$ può essere definita in modo specifico per ogni tipo di blocco ideale T , ad esempio $\varepsilon(C_i, C_j; T) = w(T)\delta(C_i, C_j; T)$. Dove $w(T) > 0$ è un peso relativo al tipo di blocco T . Nella definizione di base si ipotizza un peso uguale per tutti i tipi di blocchi \mathcal{T} . Estendendo l'inconsistenza del blocco al set di tipi di blocchi possibili \mathcal{T} si ha $\varepsilon(C_i, C_j; \mathcal{T}) = \min_{t \in \mathcal{T}} \varepsilon(C_i, C_j; T)$. Le deviazioni, inconsistenze, del blocco viene combinata in un errore totale, (*total inconsistency*), che definisce la funzione criterio del blockmodel, dove le inconsistenze dei singoli blocchi vengono sommati per tutti i blocchi:

$$P(\mathbf{C}; \mathcal{T}) = \sum_{(t,w) \in \mathbb{Z} \times \mathbb{Z}} \varepsilon(C(t), C(w); \mathcal{T}) \quad (2.2)$$

È possibile impostare delle penalità su specifici tipi di blocchi in modo che le inconsistenze con questi tipi di blocchi siano maggiormente penalizzate. Ad esempio, se i blocchi nulli sono particolarmente importanti, le inconsistenze in tali blocchi possono essere pesate specificando una penalità maggiore. Il risultato, in genere, è che i blocchi nulli empirici sono identificati senza inconsistenze portando ad una corrispondenza perfetta con il blocco nullo ideale. Queste penalità vanno a modificare il risultato della funzione criterio, che in genere, subisce un incremento e rende non confrontabili direttamente funzioni criterio definite per blockmodel strutturati diversamente. L'applicazione del Blockmodeling Generalizzato a casi reali porta frequentemente a molteplici partizioni, tutte ben adattate ai dati, con funzione criterio pari a 0. In questi casi la scelta tra la partizione più appropriata deve essere guidata da considerazioni sostantive, teoriche e contestuali (83). Altrettanto frequente è il caso in cui la funzione criterio non sia esattamente 0. In questo caso si può provare ad esaminare meglio la struttura del

network e specificare tipi di blocchi ideali che meglio la descrivono. Ovviamente, la minimizzazione della funzione criterio non è il fine in sé del Blockmodeling, ma lo è l'individuazione di una combinazione di tipi di blocchi che rifletta la struttura sottostante il network, di conseguenza vengono preferite partizioni con un valore della funzione criterio non minimo in senso assoluto ma con un modello di blockmodel significativo in termini sostantivi. Questo porta alla necessità di una scelta molto accurata e motivata dei tipi di blocchi ideali inseriti nel modello di struttura del blockmodel. Selezionati i tipi di blocchi si può procedere con la procedura del Blockmodeling che confronta il modello di struttura ideale, con i set di tipi di blocchi ideali più appropriati, e la struttura di dati empirici selezionando le partizioni che meglio riflettono la struttura originaria. Nella scelta dei tipi di blocchi ideali particolare attenzione deve essere rivolta all'uso dei blocchi *regular*, *row-regular* e *col-regular*. Questi tipi di blocchi, essendo poco vincolanti, tendono ad essere molto frequenti, pertanto è necessario specificarli qualora ci siano motivazioni concrete legate al modello di connessioni reale (83). In genere, viene suggerito di partire da blocchi ideali più restrittivi, come *row-dominant* e *col-dominant* (che implicano i *row-regular* e *col-regular*), e successivamente inserire quelli più flessibili. I tipi di blocco *row-regular* e *col-regular*, e *row-functional* e *col-functional*, possono essere necessari per network molto sparsi e con zone sparse del network. Nell'analisi di Blockmodeling generalizzato, data la possibilità di poter giungere facilmente ad una partizione ben modellata, è necessario che ogni singola scelta sia guidata da motivazioni concrete, con il ricorso alla conoscenze teoriche e/o empiriche, altrimenti il rischio è giungere ad un blockmodel ben adattato ma difficile da interpretare o incoerente. Particolare cautela va riservata anche all'utilizzo dei diagrammi del network come strumento di validazione e interpretazione delle partizioni in quanto poco affidabile (83).

Blockmodeling Pre-specificato. Un ulteriore sviluppo nell'esecuzione e nella valutazione del blockmodel può essere ottenuto attraverso la specificazione del blockmodel prima dell'analisi (11). Tale specificazione riguarda alcuni elementi quali la posizione del tipo di blocco ideale o la composizione dei cluster, e può essere realizzata sulla base delle conoscenze aggiuntive possedute dal ricercatore.

2.4 Differenti approcci al Blockmodeling

		Clustering	Blockmodel
Inductive		?	?
Deductive	Prespecification Constraints Constrained Prespecification	? given given	given ? given

Figura 2.4: Approcci Induttivo e Deduttivo al Blockmodeling (83)

Tali conoscenze possono riguardare gli attori (ad esempio, attributi demografici e ruoli) oppure la struttura dei legami tra le posizioni, ossia, ipotesi sulla struttura radicate nelle conoscenze teoriche (struttura gerarchica, *core-periphery*, ecc.). Questa opportunità mette in luce due possibili approcci nell'ambito del Blockmodeling Generalizzato. Un primo approccio viene definito come induttivo, o esplorativo, in cui non vi è alcuna specificazione poiché il ricercatore non dispone di informazioni aggiuntive ai dati relazionali, quindi sia il clustering sia il blockmodel devono emergere dall'analisi in base informazioni strutturali dei dati(83). Nell'approccio esplorativo, l'unica informazione iniziale è data dal set di tipi di blocchi ammessi dal tipo di equivalenza selezionato, quindi, nel caso si selezionino l'equivalenza strutturale tutti i *pattern* di connessioni empirici verranno confrontati solo con i tipi di blocchi ideali ammessi, ossia nullo e completo. Tale approccio, come evidenziato in precedenza, può essere adottato anche come step preliminare ad un Blockmodeling pre-specificato. Il secondo approccio viene definito come deduttivo, o confermativo, in esso gli elementi del clustering e del blockmodel vengono specificati prima dell'analisi, in parte o totalmente, sulla base di alcune conoscenze teoriche o empiriche. Tale approccio può assumere varie forme e gradi di rigidità a seconda degli elementi che vengono specificati e dei vincoli che vengono imposti alla struttura del blockmodel. La Figura 2.4, riporta nel dettaglio i tipi di Blockmodeling che possono essere realizzati.

Nel caso di un Blockmodeling pre-specificato, il set di cluster non viene specificato, ma viene specificato il blockmodel (la localizzazione di almeno un tipo di blocco). Nel caso di un Blockmodeling *constrained* (vincolato), viene forzata la struttura del clustering, specificando uno o più cluster, ma non la struttura blockmodel, oltre ai tipi di blocchi ideali consentiti. I vincoli possono assumere varie forme, ad esempio due unità possono essere specificate come appartenenti, o come non appartenenti, allo stesso cluster, oppure può essere imposto un certo

numero di unità in ciascun cluster. Infine, nel caso di un Blockmodeling vincolato e pre-specificato, viene vincolata la struttura del clustering e viene pre-specificato il blockmodel. Questo rappresenta la forma di Blockmodeling più rigida. Si ha un blockmodel vincolato completamente pre-specificato quando la struttura del clustering viene completamente impostata, e per tutti i tipi di blocco viene pre-specificata una collocazione precisa nella struttura. Si ha un blockmodel vincolato quando la struttura del clustering viene fissata ma non tutte le collocazioni dei tipi di blocchi ideali vengono specificate. Si ha un blockmodel completamente pre-specificato quando tutte le collocazioni dei tipi di blocchi ideali vengono specificate ma non la struttura del clustering. Indipendentemente dalla forma dei vincoli e dalla pre-specificazione, l'analisi attraverso il Blockmodeling generalizzato ricerca le soluzioni che si adattano perfettamente ai dati oppure quelle che si avvicinano maggiormente ad essi considerato il blockmodel specificato e i vincoli imposti. Come criterio guida per valutare il grado di adattamento ai dati del blockmodel specificato viene utilizzato il valore di un'appropriata funzione di criterio (15, 83). Infine, è necessario sottolineare che oltre ai tipi di blocchi ideali, che descrivono i modelli di connessioni all'interno e tra le posizioni, è possibile considerare una serie di tipi di blockmodel teorici, che, sulla base di specifici modelli teorici, descrive i modelli di connessioni dell'intero network. Ne sono un esempio il modello dei sottogruppi coesi, in cui le connessioni sono presenti solo all'interno delle singole posizioni; il modello *core-periphery*, in cui vi è una posizione centrale (*core*), internamente coesa, legata con tutte le altre posizioni (periferia) che, a loro volta, sono tutte legate alla posizione *core*, ma non sono connesse tra di loro e neanche coese al loro interno (90, 91). Questi sono solo due dei modelli teorici possibili, che possono essere utilizzati come struttura teorica per il blockmodel, da confrontare con la struttura empirica del network, per valutare se, e in che misura, rifletta i modelli di connessioni realmente esistenti. Nelle applicazioni empiriche è possibile osservare delle strutture complesse che possono essere descritte dalla combinazione di più modelli di strutture ideali(192) .

2.5 Recenti Sviluppi

La possibilità di definire nuovi tipi di blocchi ideali sulla base sia di considerazioni formali, legate al tipo di equivalenza, sia sostanziali, legati a specifiche motivazioni teoriche sui modelli di connessione, consente di estendere notevolmente il campo di applicazione delle procedure di Blockmodeling Generalizzato. Inoltre, la sua flessibilità consente l'analisi di strutture più complesse di quelle su cui esso è stato originariamente formulato. Tra gli sviluppi sul Blockmodeling realizzati negli ultimi anni, alcuni riprendono vecchie idee trattandole in modo nuovo. Tra quelli più rilevanti si possono individuare l'estensione del Blockmodeling per network a più vie e a più modi (*multi-way*, *multi-mode*) (17, 26, 26, 47, 49, 50, 51, 52, 81, 83, 173); Blockmodeling per network pesati (*valued*, o *weighted*, *network*) (46, 204); *stochastic* Blockmodeling (3, 122, 127); Blockmodeling per *signed* network (76, 78, 79, 84); Blockmodeling per *complex network* e *large network* (129, 170, 189); *Multilevel* Blockmodeling (206), Blockmodeling per *sparse network* (167, 205). Tra gli sviluppi sul Blockmodeling realizzati in un *framework* di Blockmodeling Generalizzato, quello che risulta particolarmente utile ai fini della presente proposta metodologica riguarda l'applicazione a network con più di un modo, nello specifico i network *two-mode*, o di affiliazione, che rappresentano la struttura di riferimento nella presente proposta. Pertanto tra i possibili sviluppi nel seguito viene presentata l'estensione del Blockmodeling Generalizzato a network *two-mode*.

2.5.1 Estensione del blockmodeling a network *two-mode*

La prima formulazione di un approccio di Blockmodeling per strutture diverse da quelle *one-mode* si deve a Borgatti e Everett (26) che ne delineano i tratti essenziali del Blockmodeling per matrici *multi-mode* e forniscono una prima strategia per realizzarlo. Elementi caratteristici di questa prima formulazione sono la necessità di partizionare simultaneamente le righe e le colonne, e la definizione di un criterio per la realizzazione di Blockmodeling sia strutturale sia regolare. Ricalcando la distinzione tra equivalenza strutturale e regolare, nel primo caso due attori sono strutturalmente equivalenti se partecipano esattamente alle stesse attività e, allo stesso tempo, due attività sono strutturalmente equivalenti se

sono frequentate esattamente dagli stessi attori. Nel secondo caso due attori sono regolarmente equivalenti se partecipano ad attività equivalenti e, allo stesso tempo, due attività sono regolarmente equivalenti se sono frequentate da attori equivalenti. Questa estensione è stata riformulata da Doreian, Batagelj e Ferligoj (81, 83) come un problema di Blockmodeling generalizzato. Come elemento di partenza nell'estensione del Blockmodeling Generalizzato al network *two-mode* vi è un cambiamento nella struttura dei dati di riferimento che non è più una matrice quadrata $n \times n$ ma una matrice rettangolare $n \times m$. Inoltre, vi è un cambiamento nel significato legato alla natura della struttura. Nei network *two-mode* non ci sono legami diretti tra i due set che la compongono, ma ognuno di questi set viene studiato in relazione ai legami che ha con l'altro set. Nel caso di un network di affiliazione, caratterizzato da un set di attori e da un set di eventi, o attività, a cui gli attori prendono parte, è possibile analizzare le relazioni esistenti tra gli attori a partire dalle partecipazioni congiunte agli eventi, ed è possibile analizzare le relazioni esistenti tra gli eventi a partire dalla comune condivisione di alcuni attori. Le caratteristiche della struttura di un network *two-mode* richiedono specifici adattamenti alle tecniche di Blockmodeling Generalizzato (81, 83). Essendo di forma rettangolare e non quadrata, come un network *one-mode*, non si può più ragionare in termini di blocchi diagonali e extra-diagonali. Inoltre, in un *one-mode* blockmodel, le righe e le colonne sono partizionate simultaneamente e allo stesso modo. Mentre, nel caso delle matrici *two-mode*, le righe e le colonne non possono essere partizionate nello stesso modo. Ossia le partizioni delle righe e colonne non sono identiche, esse, sia concettualmente sia tecnicamente, costituiscono due entità differenti. In termini formali un network *two-mode*, $\mathcal{N} = (\mathcal{V}_1, \mathcal{V}_2, R)$, è costituito da due set disgiunti di vertici, o nodi o unità, \mathcal{V}_1 e \mathcal{V}_2 , con $\mathcal{V}_1 = (a_1, a_2, \dots, a_n)$ e $\mathcal{V}_2 = (e_1, e_2, \dots, e_m)$, in cui le unità in un set possono essere connesse solo con le unità dell'altro set, $\mathcal{V}_1 \cap \mathcal{V}_2 = \emptyset$; e una relazione R che definisce i legami tra i due set di unità $R \subseteq \mathcal{V}_1 \times \mathcal{V}_2$. Le relazioni all'interno di network *two-mode* possono essere rappresentate anche attraverso una matrice rettangolare $\mathbf{F} = [f_{ij}]_{n \times m}$ dove $f_{ij} = 1$ se esiste un legame tra l'unità a_i e l'unità e_j , oppure $f_{ij} = 0$ se il legame è assente. L'approccio per il *two-mode* Blockmodeling (Generalizzato) segue gli stessi presupposti teorici generali per network *one-mode* ma è sviluppato per ottenere un two-clustering (81, 83). Sia $\mathbf{C} = (\mathbf{C}_1, \mathbf{C}_2)$ un

two-clustering dove $\mathbf{C}_1 = \{C_{1_n}, C_{2_n}, \dots, C_{K_1}\}$ è una partizione (o clustering) di \mathcal{V}_1 in K_1 cluster e $\mathbf{C}_2 = \{C_{1_m}, C_{2_m}, \dots, C_{K_2}\}$ è una partizione del set \mathcal{V}_2 in K_2 cluster (con $1 \leq K_1 \leq n$ e $1 \leq K_2 \leq m$), tale partizione con K_1 e K_2 cluster si dice (K_1, K_2) – *partition*. Sia ϕ è un set di tutti i possibili two-clustering o partizioni. L’obiettivo del *two-mode* Blockmodeling è quello di cercare di individuare nel network \mathcal{N} un two-clustering $\mathbf{C} = (\mathbf{C}_1, \mathbf{C}_2)$ tale che rispecchi i tipi di blocchi selezionati (81, 83). Anche in questo caso il problema del *two-mode* Blockmodeling generalizzato viene formulato come un problema di ottimizzazione (ϕ, P, \min) . Quindi, si vuole determinare il two-clustering $\mathbf{C}^* = (\mathbf{C}_1^*, \mathbf{C}_2^*) \in \phi$ per cui:

$$P(\mathbf{C}^*) = \min_{\mathbf{C} \in \phi} P(\mathbf{C}), \quad (2.3)$$

dove ϕ è il set di tutti i possibili two-clustering e P è la funzione criterio. La funzione criterio di $P(\mathbf{C})$ viene ottenuta nello stesso modo del *one-mode* Blockmodeling:

$$P(\mathbf{C}) = P((\mathbf{C}_1, \mathbf{C}_2)) = \sum_{C_i \in \mathbf{C}_1, C_j \in \mathbf{C}_2} \min_{B \in \mathcal{B}(C_i, C_j)} d(R(C_i, C_j), B), \quad (2.4)$$

dove $\mathcal{B}(C_i, C_j)$, con $C_i \subseteq \mathcal{V}_1$ e $C_j \subseteq \mathcal{V}_2$, indica il set di tutti i blocchi ideali corrispondenti al blocco $R(C_i, C_j)$, e l’espressione $d(R(C_i, C_j), B)$ misura la differenza (ossia in numero di inconsistenze) tra il blocco empirico $R(C_i, C_j)$ ed il corrispondente blocco ideale B . Anche in questo caso la funzione d deve essere compatibile con il tipo di equivalenza selezionato. La funzione criterio $P((\mathbf{C}_1, \mathbf{C}_2))$ è sensibile se, e solo se, $P((\mathbf{C}_1, \mathbf{C}_2)) = 0$, quindi il blockmodel determinato attraverso $(\mathbf{C}_1, \mathbf{C}_2)$ è esatto, e viceversa. Il problema di ottimizzazione così formulato viene risolto con una procedura di ottimizzazione locale (*local optimization*).

Accanto all’approccio appena esposto ne viene definito un altro, equivalente al precedente, dove il *two-mode* Blockmodeling viene gestito come un problema di Blockmodeling Generalizzato di un network *one-mode*. Il network *two-mode* viene convertito in un network *one-mode* dato dall’unione dei due set $\mathcal{N} = (\mathcal{V}_1 \cup \mathcal{V}_2, R)$, unica differenza è l’introduzione di un vincolo di bipartizione (*constraint*) $\forall C \in \mathbf{C} \in \phi : ((C \subseteq \mathcal{V}_1) \vee (C \subseteq \mathcal{V}_2))$. Tale vincolo impedisce ad un cluster C di contenere unità di entrambi i set \mathcal{V}_1 e \mathcal{V}_2 (81, 83). L’approccio

del *two-mode* Blockmodeling può essere utilizzato anche per analizzare matrici *one-mode* asimmetriche, in cui i legami in ingresso e in uscita vengono considerati come due set differenti. Anche un *two-mode* Blockmodeling può essere realizzato secondo un approccio esplorativo o confermativo, a seconda che si abbiano o meno informazioni aggiuntive di carattere teorico e/o empirico. È possibile l'aggiunta di penalità o pesi per i tipi di blocchi ideali che consentono di delineare ulteriormente la struttura del network (83). Inoltre, come nel *one-mode* Blockmodeling, le informazioni aggiuntive sulla struttura sociale esaminata possono essere utilizzate come guide nella scelta tra partizioni ben modellate. Nel caso del *two-mode* Blockmodeling, a fianco dei notevoli vantaggi legati alla possibilità di partizionare, simultaneamente, ma in modo differente, due distinte classi di entità rispettando la natura della struttura originaria, è possibile individuare alcune criticità. Viene posta in evidenza la complessità legata alla realizzazione del *two-mode* Blockmodeling, che può in parte essere mitigata dal ricorso alla pre-specificazione. Anche questa strategia, tuttavia, deve essere gestita con cautela in quanto l'utilizzo di diversi tipi di blocchi, senza una motivazione sostantiva di fondo, può incorrere nel rischio di *over-fitting* creando blockmodel con valori bassi per la funzione criterio ma poco chiare e interpretabili. Inoltre, l'utilizzo di una procedura di clustering di ottimizzazione locale, non garantisce che siano individuate tutte le partizioni possibili in grado di minimizzare la funzione criterio e, allo stesso tempo, è possibile ottenere molteplici partizioni aderenti ai dati, con una funzione criterio pari a zero. Su questi punti sono state formulate diverse proposte (46, 47, 48, 49, 51, 52). Tra i due approcci al *two-mode* Blockmodeling sino ad ora esposti, si ritiene che il primo sia quello che maggiormente risponda all'esigenza di un approccio diretto all'analisi di un network *two-mode* rispettandone la natura della specifica struttura, pertanto sarà questo l'approccio su cui verranno applicati gli sviluppi metodologici proposti in questo lavoro, con particolare attenzione agli aspetti legati al ruolo delle informazioni aggiuntive nella definizione del blockmodel.

2.6 Blockmodeling: vantaggi e problemi aperti

Numerosi sono i vantaggi che possono essere ricondotti all'utilizzo delle tecniche di Blockmodeling, soprattutto nel caso del Blockmodeling Generalizzato (75, 77, 83, 98). Alcuni tra i più rilevanti sono 2.4:

- (i) Utilizzo di una funzione criterio esplicita, compatibile con un certo tipo di equivalenza, che consente l'integrazione all'interno del blockmodel di una misura di bontà del grado di adattamento ai dati reali.
- (ii) Introduzione di un ampio numero di tipi di blocchi, oltre a quelli nullo, completo e regolare.
- (iii) Possibilità di costruire nuovi tipi di blocchi e con essi nuovi blockmodel su basi sostantive teoricamente radicate, agevolando l'operativizzazione di concetti astratti.
- (iv) Approccio deduttivo, e non solo induttivo. La pre-specificazione del blockmodel consente non solo la specificazione dei tipi di blocchi ma anche la loro collocazione nella struttura del blockmodel.
- (v) Possibilità di inserire vincoli nella struttura del clustering e impostare penalità su specifici tipi di blocchi in modo che le inconsistenze con questi tipi di blocchi siano maggiormente penalizzate.

Accanto ai vantaggi appena illustrati, nei più recenti sviluppi sul Blockmodeling Generalizzato, possono essere individuati alcuni problemi aperti che lasciano spazio ad ulteriori approfondimenti (75, 77, 98) :

- (i) Specificazione dei confini del network (rilevanza/irrelevanza delle unità incluse) e conseguenze sulla vulnerabilità del Blockmodeling Generalizzato e sulla validità del blockmodel stabilito empiricamente.
- (ii) Sensibilità alle caratteristiche dei dati (es. dati mancanti e inaccurati).

- (iii) Definizione di una rappresentazione più ampia dell'ambiente in cui il network è inserito in modo da preservare la validità del blockmodel stabilito anche se i dettagli dei collegamenti con il resto del network globale non sono inclusi nei dati.
- (iv) Trattamento dei network di grandi dimensioni (large network);
- (v) Capacità di identificare la struttura fondamentale nel network nonostante le informazioni siano incomplete.
- (vi) Estensione del Blockmodeling per strutture di dati non binarie e a più di una classe di entità (es. network pesati, network *multi-mode*).
- (vii) Valutazione statistica della bontà del blockmodel stabilito empiricamente e gestione delle inconsistenze (come definire un limite oltre il quale il blockmodel non è più valido).
- (viii) Gestione e trattamento delle relazioni multiple;
- (ix) Blockmodeling dinamico per tenere conto dell'evoluzione del network;
- (x) Criteri per l'individuazione del numero di posizioni nel blockmodel.

In aggiunta a quelli sopra indicati è possibile individuare alcune problematiche relative all'applicazione del Blockmodeling al network *two-mode*, che costituisce la struttura di riferimento negli sviluppi della presente proposta metodologica.

- (i) Dimensioni elevate del network che comportano soluzioni più instabili con più partizioni plausibili (*well fitting*), elevati valori della funzione criterio e difficile identificazione sia del numero delle posizioni sia dei componenti al loro interno, con un incremento del carico computazionale combinatorio(77, 81, 83).
- (ii) Interpretazione e spiegazione dei "fenomeni di network". Utilizzo privilegiato di dati relazionali e proprietà strutturali (misure relative alle caratteristiche del network) a discapito degli attributi descrittivi delle entità(113).

- (iii) Difficoltà di operare una scelta tra partizioni ugualmente ben adattate senza ulteriori informazioni, sia teoriche sia empiriche, sul fenomeno sociale in esame.
- (iv) Necessità di informazioni sostantive o empiriche che orientino le scelte legate alla pre-specificazione (81, 83).
- (v) Visualizzazione dei dati relazionali. I risultati del *two-mode* Blockmodeling possono essere rappresentati mediante grafo ridotto (grafo immagine), ma senza la possibilità di valutare la distanza tra le posizioni.

2.7 Proposta metodologica

Come si è visto, uno degli sforzi conoscitivi più rilevanti nella SNA è legato allo sviluppo di metodi che consentano l'analisi e la comprensione della struttura dei modelli di connessioni alla base di un network sociale e delle sue conseguenze. L'analisi dei modelli di connessioni sono fondamentali per la comprensione delle azioni individuali e del funzionamento del sistema sociale nel suo complesso. All'interno di queste elaborazioni assumono particolare rilievo i metodi che consentono l'analisi delle strutture di dati *two-mode* (o *multi-mode*) che racchiudono la dualità della struttura sociale (43), ossia la dinamica individuo/collettività espressa attraverso i legami di appartenenza, o affiliazione, che connettono gli attori a differenti livelli della struttura sociale (individuo/ gruppo/ organizzazione/ comunità).

Le tecniche di Analisi Fattoriali, in particolare l'Analisi delle Corrispondenze, e le tecniche di Blockmodeling sono tra gli strumenti più appropriati per sintetizzare, analizzare e rappresentare graficamente sia la struttura sottostante il network sia i sottoinsiemi significativi che compongono un network *two-mode*, rispettandone la natura duale.

La seguente proposta si presenta come sviluppo di una strategia che consenta di affrontare alcune delle problematicità legate all'applicazione delle tecniche di Blockmodeling per l'analisi dei network *two-mode*. Tale strategia prevede il ricorso all'ausilio degli strumenti offerti dalle tecniche di Analisi Fattoriale.

L'approccio sviluppato è vicino, ma non completamente incorporato, in quello

che è stato definito come approccio posizionale. In esso, l'obiettivo principale è quello di svelare la struttura sottostante il network identificando sottoinsiemi di unità e di relazioni rilevanti nella struttura relazionale, quali "posizioni e ruoli". In questo contesto, l'approccio adottato cerca di combinare i vantaggi principali sia delle tecniche di Analisi Fattoriale sia di quelle Posizionale che, pur provenendo da tradizioni matematiche e statistiche differenti, presentano interessanti punti di connessione. Entrambe cercano di individuare e sintetizzare la struttura sottostante i fenomeni osservati guardandoli da una prospettiva globale. Entrambi i metodi sono considerati compatibili e complementari in quanto considerano lo stesso oggetto con diversi livelli di dettaglio e da diversi punti di vista, consentendone un'analisi approfondita.

Seguendo un approccio diretto, si sceglie di analizzare e rappresentare i network *two-mode* attraverso l'Analisi delle Corrispondenze Multiple (ACM) con codifica disgiuntiva completa adottando la prospettiva del *Doubling* (70, 71, 119). La scelta di questa versione dell'ACM è legata al fatto che essa: (i) è stata sviluppata per il trattamento di matrici casi-per-variabili, che hanno una struttura simile ai network *two-mode*; (ii) assegna un ruolo diverso alle due classi di entità, corrispondenti agli attori e agli eventi nel network *two-mode*, pertanto, consente di evidenziare le caratteristiche distintive in ogni modo; (iii) permette di aggiungere nell'analisi variabili e punti supplementari al fine di migliorare l'interpretazione dei risultati; e (iv) nella prospettiva del *Doubling* consente di realizzare delle visualizzazioni interpretabili in termini di somiglianze tra le relazioni degli attori/eventi nel network che possono essere lette non solo in termini di co-partecipazione ma anche di "non" co-partecipazione. Questi ultimi due vantaggi offrono la possibilità di inserire i risultati del Blockmodeling come variabile supplementare e di proiettarli in uno spazio metrico. Ciò consente di valutare la distanza tra le posizioni e caratterizzarle grazie al contributo delle informazioni aggiuntive, quali attributi descrittivi e proprietà strutturali delle unità.

Nell'analisi dei network di una certa dimensioni, e sparsi, l'ACM può essere di supporto al *two-mode* Blockmodeling secondo tre principali direzioni. Nell'ambito del *two-mode* Blockmodeling Esplorativo si propone di applicare il metodo di decomposizione di Eckart-Young (85) ed eseguire il Blockmodeling su una matrice di affiliazione di rango inferiore. Questa operazione consente di estrarre la

struttura relazionale "forte", eliminando in parte le inconsistenze, e facilita l'operazione di partizione. La scelta di tale strategia come fase esclusiva è da valutare con cautela in quanto nelle struttura "debole", che viene scartata insieme alle inconsistenze, potrebbero celarsi *pattern* relazionali di un certo interesse. Tale operazione di riduzione può essere suggerita come fase preliminare al Blockmodeling sul network originale e consente di ottenere una visione globale della struttura e di acquisire informazioni utili come guida per alcune scelte iniziali quali il numero di cluster. L'ACM può essere di supporto al *two-mode* Blockmodeling Pre-specificato in quanto, insieme al numero di cluster come per l'altra direzione, può fornire ulteriori informazioni da utilizzare nella specificazione modello di struttura del blockmodel teorico che si vuole testare. La pre-specificazione del modello richiede delle informazioni aggiuntive, a quelle relazionali, di carattere teorico o empirico. L'ACM, consentendo di introdurre di informazioni aggiuntive risulta molto efficace a tale scopo. Quest'ultima fornisce un utile supporto alla ricostruzione del contesto della struttura relazionale e alla definizione dei possibili vincoli (*constraint*) da inserire nel modello, con conseguenze positive sulla stabilità delle soluzioni finali. In aggiunta, e in modo non del tutto separato rispetto alle direzioni precedenti, l'ACM offre la possibilità di integrare al suo interno, come variabili e punti supplementari, gli stessi risultati del Blockmodeling. Di conseguenza, il blockmodel può essere messo in relazione alla struttura dei legami sottostanti, e alle altre informazioni descrittive e strutturali, sia dal punto di vista analitico sia visuale. In effetti, uno dei vantaggi più interessanti è dato dal fatto che l'ACM consente la proiezione delle posizioni in uno spazio metrico che rende possibile una rappresentazione pienamente significativa, in cui lo spazio tra le posizioni può essere interpretato in termini di distanze. Un vantaggio ulteriore è dato dalla possibilità di ottenere informazioni analitiche utili ad orientare la scelta tra partizioni egualmente ben modellate (*equally well fitting partition*). In questo modo, partizioni analoghe, rispetto ad un certo criterio di equivalenza, possono essere confrontate non solo in base a considerazioni di carattere teorico ma anche empirico. Inoltre, l'opportunità di inserire le posizioni in un contesto informativo più ampio ne consente la caratterizzazione e una interpretazione più profonda. Dal canto suo, il Blockmodeling fornisce una partizione simultanea, ma distinta, delle righe e delle colonne della matrice *two-mode*, consentendo di

mettere in relazione diretta le unità di un set con quelle dell'altro set dalle quali sono definite. Quindi, le classi di attori equivalenti, ossia le posizioni, vengono definite a partire dai loro modelli di partecipazione ad eventi, o appartenenza a gruppi, che, a loro volta, costituiscono classi di eventi, o gruppi, equivalenti in funzione dei modelli di partecipazione degli attori che ne fanno parte. Inoltre, le rappresentazione ridotte ottenute tramite il Blockmodeling, ossia il grafo ridotto e la matrice immagine, consentono di visualizzare in modo sintetico le posizioni e le relazioni tra le posizioni (blocchi). Pertanto, l'integrazione di entrambe le tecniche offre l'opportunità di valorizzare i singoli contributi informativi e migliorare l'interpretazione globale dei risultati con riflessi positivi sulla capacità di comprensione del fenomeno sociale in esame. Con tali premesse, la presente proposta può essere considerata come un approccio *multi-step* in cui le due tecniche possono interagire in una o più fasi del processo di analisi dei dati, a seconda delle specifiche esigenze.

Le opportunità derivanti dalla combinazione di tecniche di Analisi Fattoriale e tecniche di Blockmodeling sono visibili anche in una prospettiva più ampia legata alla gestione di reti eterogenee. I fenomeni sociali sono complessi, multidimensionali e pertanto la loro comprensione richiede la capacità di mettere in relazione aspetti diversi a differenti livelli di analisi. Ne consegue la necessità di trattare informazioni molteplici e di natura diversa, spesso, in un'ottica diacronica.

Nel tentativo di compiere un passo in questa direzione si propone una strategia di analisi che combini l'Analisi Fattoriale Multipla (AFM) (87, 169) e il Blockmodeling per network *two-mode* eterogenei, in cui sono presenti dati relativi ai legami tra due classi di entità differenti, in momenti temporali diversi, con attributi descritti delle entità e i modelli di relazioni che intercorrono tra loro. Questa strategia, pur non essendo in grado di cogliere pienamente la dinamica processuale alla base dell'evoluzione della struttura, consente di esplorare, descrivere e rappresentarne in modo accurato gli esiti mettendoli in relazione reciproca.

Tale approccio analitico sarà adottato per lo studio del "campo della produzione teatrale" nell'ambito del sistema regionale campano (182). Nello specifico, lo studio è volto a cogliere la struttura del campo delle co-produzioni teatrali attraverso una integrazione della Teoria del Campo di Pierre Bourdieu (35, 36) e dell'Approccio Posizionale (54, 56, 83) della SNA. Tale scelta nel tentativo di evidenziare

come l'approccio multidimensionale, privilegiato da Bourdieu per il suo carattere "relazionale" capace di portare alla luce le "relazioni oggettive", e quello della SNA, "scartato" dallo stesso per essere eccessivamente ingessato nella descrizione di particolari legami piuttosto che alla struttura sottostante, possano essere non solo compatibili ma anche fruttuosamente integrabili. Il punto di partenza della connessione tra "struttura del network" e "struttura oggettiva" è legato al fatto che entrambe sono considerate come parte del campo della produzione culturale. Si ritiene che l'approccio posizionale della SNA sia in grado di rivelare la struttura di relazioni che sostiene quelle oggettive e che, contemporaneamente, costituisce il *frame* in cui essa prende forma.

Capitolo 3

Blockmodeling e Analisi Fattoriale per Network *Two-mode*

3.1 Analisi Fattoriale e Blockmodeling: legami

La presente proposta si basa sull'utilizzo combinato delle tecniche di Analisi Fattoriale e delle tecniche di Blockmodeling per l'analisi dei network *two-mode*. Come si è detto, tale approccio è vicino, ma non completamente incorporato nell'approccio posizionale. In esso, l'obiettivo principale è quello di svelare la struttura sottostante il network identificando sottoinsiemi di unità e di relazioni rilevanti nella struttura relazionale, quali "posizioni e ruoli". In questo contesto, l'approccio adottato cerca di combinare i vantaggi principali sia delle tecniche di Analisi Fattoriale sia di quelle Posizionale che, pur provenendo da tradizioni matematiche e statistiche differenti, come si vedrà nel corso del capitolo, presentano interessanti punti di connessione. Entrambe cercano di individuare e sintetizzare la struttura sottostante i fenomeni osservati guardandoli da una prospettiva globale. Nello specifico è possibile rintracciare una forte connessione tra l'Analisi delle Corrispondenze Multiple (ACM) e il Blockmodeling poichè entrambe consentono di ricostruire le strutture di similarità nel network sulla base di un certo criterio di equivalenza tra le unità. Come si vedrà, l'ACM, ma anche l'Analisi delle Corrispondenze (AC), si basano sul principio di equivalenza distribuzionale, mentre il Blockmodeling si basa sul principio di equivalenza strutturale, almeno nella sua forma originaria. Queste due forme di equivalenza hanno una natura

3.2 Caratteristiche dell'Analisi delle Corrispondenze Multiple per network *two-mode*

diversa. Nel primo caso è legata alle distribuzioni di frequenza dei profili di riga e di colonna, mentre, nel secondo caso alle relazioni tra le unità. Tuttavia, esiste una forte connessione tra questi due principi. Tale connessione diventa più forte ed evidente nel caso dell'utilizzo dell'ACM, poichè il complesso sistema di pesi che ne è alla base, legato alla metrica del χ^2 , rende le distanze che ne risultano molto simili alla distanza Euclidea δ_E^2 ponderata. A sua volta la distanza Euclidea è una misura sensibile e compatibile con l'equivalenza strutturale. Di conseguenza, le rappresentazione grafiche realizzate attraverso l'ACM sono basate su proprietà che gli consentono di riprodurre la stessa struttura relazionale emergente dal Blockmodeling basato sull'equivalenza strutturale. Si ritiene che tale connessione legittimi la scelta di combinare le due tecniche analizzando le opportunità derivanti da uno scambio reciproco. Di seguito verranno esaminate in dettaglio le relazioni tra l'ACM e il Blockmodeling nell'ambito dell'analisi dei network *two-mode*.

3.2 Caratteristiche dell'Analisi delle Corrispondenze Multiple per network *two-mode*

L'interesse nell'analisi delle strutture *two-mode*, in special modo quelle di affiliazione, è in continuo incremento nell'ambito delle scienze sociali, ma non solo, (1). Come si è visto nei capitoli precedenti, un network di affiliazione è un tipo di network *two-mode* caratterizzati da un set di attori e un set di eventi (o gruppi) in cui essi sono coinvolti. Tali strutture relazionali si prestano a modellare efficacemente varie forme di "dualità", come quella persone-gruppi, partecipanti-eventi, pratiche-simboli, ecc. (43). Numerosi sono le applicazioni dei network di affiliazione in vari contesti di ricerca per studiare le cariche incrociate nei consigli di amministrazione (*interlocking corporate directorates*) (55, 109), collaborazione accademica (154), comunicazione politica (63), il linguaggio umano (202), partecipazione ad eventi sociali (43, 66), la struttura disciplinare di organizzazioni accademiche (6, 139). La struttura del network di affiliazione si differenzia da quella del network *one-mode* più diffusi nell'ambito della SNA. Quest'ultimo include le relazioni dirette tra coppie di attori appartenenti ad un'unica classe di entità. Diversamente, in un network di affiliazione non sono presenti i legami

3.2 Caratteristiche dell'Analisi delle Corrispondenze Multiple per network *two-mode*

diretti tra coppie di attori, ma sono definiti indirettamente attraverso la comune partecipazione, o appartenenza, ad un evento, o gruppo. Similmente i legami tra una coppia di eventi, o gruppi, sono definiti dalla condivisione di attori che partecipano ad entrambi. Quindi, un network di affiliazione presenta una triplice natura relazionale: (i) incorpora le relazioni dirette tra attori ed eventi; (ii) coglie i legami tra attori generati dalla comune partecipazione (*co-membership*, o *co-attendance actors*); (iii) esprime i legami tra gli eventi (o gruppi) creati dalla condivisione di partecipanti comuni (*overlapping* o *interlocking events*) (192). Queste caratteristiche dei network di affiliazione, nel presente lavoro, sono di notevole interesse. In quest'ottica, infatti, sottogruppi o classi di attori con caratteristiche strutturali comuni (equivalenti), come le "posizioni", possono essere definiti anche dalla partecipazione ad eventi comuni o esperienze (43). In questo modo le posizioni possono collocare gli attori all'interno dello spazio sociale definito dalla partecipazione ad eventi (99). Gli eventi possono servire come mezzo attraverso cui gli attori esprimono caratteristiche comuni preesistenti oppure possono servire come occasione di contatto e socializzazione per un insieme di attori favorendone le opportunità di interazione (99, 180). In questo modo gli eventi possono servire come punti focali, "foci" (97), attraverso i quali prendono forma e si manifestano i fenomeni di gruppo (97).

Nell'analisi dei network *two-mode* possono essere distinti due approcci principali: l'approccio indiretto (*conversion approach*) e l'approccio diretto (*direct approach*) (24, 27, 30, 91). Nel primo approccio, un network *two-mode* viene convertito in due network *one-mode* e l'analisi viene condotta su un modo per volta, sebbene questa strategia possa comportare la perdita di alcune informazioni (27). L'approccio diretto, che viene adoperato in questa proposta, invece, considera un network *two-mode* nella sua interezza, e i due modi vengono analizzati congiuntamente. Nell'analisi di network *two-mode*, una delle tematiche rilevanti è legata alla visualizzazione e all'analisi grafica della struttura relazionale sottostante. In quest'ottica l'approccio diretto ricorre principalmente a rappresentazioni attraverso i Grafi Bipartiti e lo Spring Embedding (30) o a Metodi Fattoriali per dati qualitativi come l'AC (20, 119). Quest'ultima consente di rappresentare, in spazi di dimensionalità ridotta, i punti corrispondenti agli attori e gli eventi al fine di valutarne la similarità nei modelli di partecipazione (30, 44, 94, 171, 191).

L'AC può essere applicata se si considera la matrice di affiliazione, corrispondente ad un network *two-mode*, come una tabella di contingenza a due vie, o come una matrice a due vie caso-per-variabile. Il primo approccio è stato adottato spesso e si traduce in un'analisi delle corrispondenze semplici della matrice affiliazione(30, 171). Nel presente lavoro, viene adottato il secondo approccio, considerando la matrice di affiliazione come una matrice rettangolare caso-per-variabile. In particolare, si adotta l'approccio dell'ACM, un'estensione dell'AC per l'analisi di molteplici variabili categoriali (118) considerato come più appropriato a questa tipo di struttura di dati (71). Seguendo l'approccio diretto, si sceglie di analizzare e rappresentare i network *two-mode* attraverso l'ACM con codifica disgiuntiva completa seguendo la prospettiva del *doubling* di Greenacre(119). La scelta di questa versione dell'ACM è motivata dal fatto che (70, 71) : (i) l'ACM è stata sviluppata per il trattamento di matrici casi-per-variabili, che hanno una struttura simile ai network *two-mode*; (ii) l'ACM assegna un ruolo diverso alle due classi di entità, corrispondenti agli attori e agli eventi nei network *two-mode*, pertanto, consente di evidenziare le caratteristiche distintive in ogni modalità; (iii) l'ACM permette di aggiungere all'analisi variabili e punti supplementari al fine di migliorare l'interpretazione dei risultati; e (iv) l'ACM nella prospettiva del *doubling* consente di realizzare delle visualizzazioni interpretabili in termini di somiglianze tra le relazioni di rete tra attori/eventi che possono essere lette non solo in termini di co-partecipazione ma anche di 'non' co-partecipazione. Questi ultimi due vantaggi offrono la possibilità di inserire i risultati del Blockmodeling come variabile supplementare e di proiettarli in uno spazio metrico. Ciò consente di valutare la distanza tra le posizioni e caratterizzarle grazie al contributo delle informazioni aggiuntive, quali attributi descrittivi e proprietà strutturali delle unità.

3.3 Analisi delle Corrispondenze Multiple in prospettiva *doubling*

Sia \mathcal{G} un network di affiliazione, composto da due set di unità connesse tra loro, attori ed eventi. Tale network può essere rappresentato come $\mathcal{G}(\mathcal{V}_1, \mathcal{V}_2, R)$ con due set disgiunti di nodi, \mathcal{V}_1 e \mathcal{V}_2 di cardinalità n e m , e un set di legami o

3.3 Analisi delle Corrispondenze Multiple in prospettiva *doubling*

archi, $R \subseteq \mathcal{V}_1 \times \mathcal{V}_2$. Per definizione $\mathcal{V}_1 \cap \mathcal{V}_2 = \emptyset$; il set $\mathcal{V}_1 = \{a_1, a_2, \dots, a_n\}$ rappresenta il set di n attori, mentre $\mathcal{V}_2 = \{e_1, e_2, \dots, e_m\}$ rappresenta il set di m eventi. Il legame $r_{ij} = (a_i, e_j)$, $r_{ij} \in R$, è una coppia ordinata di nodi, e indica se un attore a_i partecipa ad un evento e_j . Il set $\mathcal{V}_1 \times \mathcal{V}_2$ può essere pienamente rappresentato attraverso una matrice binaria *matrice di affiliazione* $\mathbf{F} = (f_{ij})$, $i = 1, \dots, n$, $j = 1, \dots, m$, con $f_{ij} = 1$ se $(a_i, e_j) \in R$ e 0 altrimenti. Dato \mathbf{F} , i marginali di riga e di colonna $f_{i\cdot} = \sum_{j=1}^m f_{ij}$ e $f_{\cdot j} = \sum_{i=1}^n f_{ij}$ coincidono con il grado, o *degree*, d_i dell' i -esimo attore e la dimensione, *size*, s_j dell' j -esimo evento, rispettivamente, $f_{i\cdot} = d_i$ and $f_{\cdot j} = s_j$. Uno dei primi tentativi di applicare l'ACM ad una struttura *two-mode* è ritracciabile nel lavoro di Bourdieu (36) che la utilizza per analizzare Università, affiliazioni professionali e le variabili legate alla formazione dei Professori parigini. Egli, tuttavia, non ha trattato la sua matrice di dati come matrice di affiliazione poiché considerava le affiliazioni dei professori come proprietà degli individui e non come relazioni(68) . Successivamente, nella SNA, l'analisi dei network di affiliazione (94, 191) è stata condotta attraverso una particolare versione di ACM (57). In questo caso, l'algoritmo dell'AC) è stato applicato ad una matrice legami per attori ed eventi, in cui i legami $r_{ij} = (a_i, e_j)$ sono assunti come unità di analisi e la matrice di affiliazione $\mathbf{F}(n \times m)$ viene trasformata in un matrice di variabili indicatrici multipla (*indicator matrix*) con $n + m$ colonne (uno di n per ogni attore e uno di m per ogni evento) e tante righe quanti sono i legami L (94). Tale approccio, tuttavia, non ha ricevuto molto consenso nella comunità scientifica(58, 59, 120). Nella presente proposta, gli attori vengono considerati come unità di analisi e la partecipazione ad eventi come variabili categoriali dicotomiche (70, 71). La matrice \mathbf{F} quindi è una matrice dati casi-per-variabili multivariata (116), in cui alle righe e alle colonne è assegnato uno status differente, in linea con la prospettiva della dualità(43, 94) . L'ACM viene eseguita applicando l'algoritmo dell'AC, basato sulla Decomposizione ai valori singolari (SVD) della matrice dei profili *doubled* normalizzata e centrata, alla *matrice delle variabili indicatrici*, \mathbf{Z} derivata da \mathbf{F} attraverso una codifica disgiuntiva completa.

Nell'ACM per i dati numerici, ogni riga di \mathbf{Z} rappresenta una osservazione e ogni colonna rappresenta una modalità di ciascuna variabile categoriale, con $z_{ij} = 1$ se la j -esima modalità della q -esima variabile categorica è presente nell'

3.3 Analisi delle Corrispondenze Multiple in prospettiva *doubling*

i -esimo caso, 0 in caso contrario. Nel caso di Q variabili categoriali la matrice delle variabili indicatrici è data da $\mathbf{Z} = [\mathbf{Z}_1 \dots \mathbf{Z}_Q]$. Se la q -esima variabile ha J_q categorie, Z_q è $I \times J_q$, e $J = \sum_{q=1}^Q J_q$ è il numero totale di categorie. In questo caso, i marginali di riga sono costanti e pari a Q , mentre i marginali di colonna corrispondono alle frequenze della categoria. Nel caso della matrice di affiliazione \mathbf{F} , al fine di applicare una codifica disgiuntiva completa, si considera ogni evento e_j come una variabile dicotomica con le categorie e_j^+ e e_j^- . Nella codifica disgiuntiva completa, ogni categoria viene descritta da una variabile *dummy*, quindi e_j^+ è una variabile *dummy* che codifica la partecipazione all'evento, e e_j^- è una variabile *dummy* codifica la "non" partecipazione. \mathbf{Z} conterrà due colonne ortogonali per ogni e_j . La matrice \mathbf{Z} è una matrice $n \times 2m$ di forma: $\mathbf{Z} \equiv [\mathbf{F}^+, \mathbf{F}^-]$, dove $\mathbf{F}^+ = (e_j^+) = \mathbf{F}$, e $\mathbf{F}^- = (e_j^-) = \mathbf{1} - \mathbf{F}^+ = \mathbf{1} - \mathbf{F}$, dove $\mathbf{1}$ è una matrice $n \times m$ tutta di 1.

La matrice delle variabili indicatrici \mathbf{Z} si trasforma in una matrice *doubled*, e l'ACM è equivalente alla AC della matrice *doubled*. Questo permette di interpretare i risultati in una prospettiva *doubling*.

I marginali di riga z_i della matrice \mathbf{Z} sono costanti ed uguali al numero di eventi m , mentre i marginali di colonna z_j sono uguali alla *size* degli eventi s_j , quando sono associati a e_j^+ , o a $n - s_j$, quando sono associati a e_j^- .

La matrice dei profili \mathbf{P} e le matrici dei pesi \mathbf{D}_a e \mathbf{D}_e adeguati per l'algoritmo dell'AC e coinvolti nella normalizzazione sono:

$$\mathbf{P} = \frac{\mathbf{Z}}{nm}, \quad \text{con} \quad nm = \sum_{i=1}^n \sum_{j=1}^m z_{ij}, \quad (3.1)$$

$$\mathbf{D}_a = \text{diag}\left(\frac{1}{n}, \dots, \frac{1}{n}\right), \quad (3.2)$$

$$\mathbf{D}_e = \text{diag}\left(\frac{s_1}{nm}, \dots, \frac{s_{2m}}{nm}\right). \quad (3.3)$$

Data \mathbf{F} , sia $\mathbf{P} = (p_{ij})$, per $i = 1, \dots, n$ e $j = 1, \dots, m$, la normalizzata di \mathbf{F} con $p_{ij} = f_{ij}/L$, dove $L = \sum_{i=1}^n \sum_{j=1}^m f_{ij}$ è il numero totale di legami; $\mathbf{D}_a = \text{diag}(d_1/L, \dots, d_n/L)$ e $\mathbf{D}_e = \text{diag}(s_1/L, \dots, s_m/L)$ sono rispettivamente, le matrici diagonali dei *degree* relativi degli attori d_i/L e le *size* relative degli eventi s_j/L .

3.3 Analisi delle Corrispondenze Multiple in prospettiva *doubling*

La matrice *doubled* centrata e normalizzata di \mathbf{Z} è data dalla matrice \mathbf{S}

$$\mathbf{S} = \mathbf{D}_a^{-1/2} \left(\frac{\mathbf{Z}}{nm} - \mathbf{D}_a \mathbf{1} \mathbf{1}^T \mathbf{D}_e \right) \mathbf{D}_e^{-1/2} = \sqrt{n} \left(\frac{\mathbf{Z}}{nm} - \frac{1}{n} \mathbf{1} \mathbf{1}^T \mathbf{D}_e \right) \mathbf{D}_e^{-1/2},$$

Dove $(1/n)\mathbf{1}$ è il vettore dei pesi degli attori e $\mathbf{1}^T \mathbf{D}_e$ è il vettore dei pesi degli eventi.

La Decomposizione ai valori singolari, SVD, di \mathbf{S} porta a:

$$\mathbf{S} = \mathbf{U} \mathbf{\Sigma} \mathbf{V}^T$$

Dove $\mathbf{\Sigma}$ è la matrice diagonale dei valori singolari, e \mathbf{U} , \mathbf{V} sono, rispettivamente, le matrici dei vettori singolari di sinistra e di destra.

Le coordinate principali per le righe e per le categorie delle colonne, sono, rispettivamente, definite come:

$$\mathbf{\Phi} = \mathbf{D}_a^{-1/2} \mathbf{U} \mathbf{\Sigma} = \sqrt{n} \mathbf{U} \mathbf{\Sigma} \quad (3.4)$$

$$\mathbf{\Psi} = \mathbf{D}_e^{-1/2} \mathbf{V} \mathbf{\Sigma} \quad (3.5)$$

e le coordinate standard per le righe e per le categorie delle colonne, sono, rispettivamente, definite come:

$$\mathbf{\Gamma} = \sqrt{n} \mathbf{U} \quad (3.6)$$

$$\mathbf{\Delta} = \mathbf{D}_e^{-1/2} \mathbf{V} \quad (3.7)$$

In genere, le prime due colonne di $\mathbf{\Phi}$ e $\mathbf{\Psi}$, o di $\mathbf{\Gamma}$ e $\mathbf{\Delta}$, vengono utilizzate per costruire le mappe bidimensionali in cui sono rappresentati graficamente i dati. Per comprendere meglio i vantaggi legati alla rappresentazione grafica dell'ACM è utile analizzare la geometria sottostante il metodo e le distanze tra i punti delle righe e delle colonne negli spazi originari e nelle mappe bidimensionali delle proiezioni (71).

Negli spazi originari, le distanze tra i profili dell'attore sono date da:

$$\mathbf{p}_{a_i} = [z_{i1}/m, \dots, z_{i2m}/m], \quad (3.8)$$

e tra i profili dell'evento sono date da:

$$\mathbf{p}_{e_{j+}} = [z_{1j}/s_j, \dots, z_{nj}/s_j] \quad \forall e_{j+} \quad (3.9)$$

$$\mathbf{p}_{e_{j-}} = [z_{1j}/(n-s_j), \dots, z_{nj}/(n-s_j)] \quad \forall e_{j-} \quad (3.10)$$

queste distanze nella metrica χ^2 , dell'AC, sono dati, rispettivamente, da:

$$\begin{aligned} \delta_{ACM}^2(\mathbf{p}_{a_i}, \mathbf{p}_{a_{i'}}) &= (\mathbf{p}_{a_i} - \mathbf{p}_{a_{i'}})^T \mathbf{D}_e^{-1} (\mathbf{p}_{a_i} - \mathbf{p}_{a_{i'}}) = \sum_{j=1}^{2m} \left(\frac{z_{ij}}{m} - \frac{z_{i'j}}{m} \right)^2 \frac{nm}{s_j} = \\ &= \frac{n}{m} \sum_{j=1}^m \frac{n}{s_j(n-s_j)} (f_{ij} - f_{i'j})^2 \end{aligned} \quad (3.11)$$

$$\begin{aligned} \delta_{ACM}^2(\mathbf{p}_{e_j}, \mathbf{p}_{e_{j'}}) &= (\mathbf{p}_{e_j} - \mathbf{p}_{e_{j'}})^T \mathbf{D}_a^{-1} (\mathbf{p}_{e_j} - \mathbf{p}_{e_{j'}}) = n \sum_{i=1}^n \left(\frac{z_{ij}}{s_j} - \frac{z_{ij'}}{s_{j'}} \right)^2 = \\ &= \frac{2n}{s_j s_{j'}} \sum_{i=1}^n [z_{ij} \neq z_{ij'}] \end{aligned} \quad (3.12)$$

dove $[\cdot]$ è l'Iverson bracket tale che:

$$[P] = \begin{cases} 1 & \text{se } P \text{ è vero;} \\ 0 & \text{altrimenti.} \end{cases}$$

Particolarmente importanti sono le implicazioni legate all'interpretazione delle distanze ottenute attraverso la metrica del χ^2 . Infatti, la distanza tra due profili di attori è uguale a zero quando gli attori partecipano agli stessi eventi. La distanza è maggiore di zero quando i modelli di partecipazione differiscono. La distanza tra due profili di evento è pari a zero quando gli eventi sono frequentati dagli stessi attori. La distanza è maggiore di zero quando i modelli di partecipazione sono diversi.

Un caso particolare di tale distanza si verifica quando si confrontano gli eventi e_j^+ e e_j^- . In questo caso, l'Eq. (3.12) diventa:

$$\delta_{ACM}^2(\mathbf{p}_{e_j^+}, \mathbf{p}_{e_j^-}) = \frac{n^2}{s_j(n-s_j)} \quad (3.13)$$

Questo implica che la distanza tra e_j^+ e e_j^- è inversamente proporzionale alle dimensioni dell'evento. Questa distanza raggiunge il suo valore minimo quando $s_j = n - s_j$, ossia quando la metà degli attori partecipare all'evento e_j e la metà non partecipano. Tale proprietà risulta di notevole interesse poichè uno degli aspetti critici nell'analisi delle matrici di affiliazione è legato alla necessità di gestire l'effetto delle differenze nella dimensione degli eventi a cui gli individui sono affiliati (30). La distanza del χ^2 rispetta il principio di equivalenza distribuzionale (20). Due profili distributivamente equivalenti corrisponderanno a due punti coincidenti nello spazio originario, e viceversa. Inoltre, quando i profili non sono distributivamente equivalenti le loro distanze χ^2 sono maggiori di zero. Nelle mappe bidimensionali ricavate attraverso le prime due coordinate principali, le distanze tra attori e tra gli eventi approssimano in modo ottimale quelle negli spazi originari (118). Quindi l'analisi nel piano fattoriale associato fornisce informazioni rispetto alle posizioni relative dei profili degli attori e degli eventi (71).

Sfruttando la caratteristica non diadica dei network *two-mode* di affiliazione, è possibile focalizzare l'analisi sia sugli attori rappresentati come punti in uno spazio (spazio degli eventi) attraversato dagli assi principali con coordinate Φ (Eq. 3.4) sia sugli eventi rappresentati come vettori in uno spazio (spazio degli attori) attraversati dagli assi principali con coordinate Ψ (Eq. 3.5). In particolare, grazie all'adozione del principio del *doubling*, la rappresentazione dell'evento nello spazio degli attori consente particolari caratteristiche ed interpretazioni (71). Inoltre, grazie alla relazione analitica tra lo spazio degli attori e lo spazio degli eventi, è possibile eseguire una rappresentazione congiunta degli attori e degli eventi in uno spazio comune per analizzare le relazioni tra i due modi del network. In aggiunta, utilizzando le modalità degli attributi come informazioni supplementari, l'ACM permette di aggiungere nell'analisi gli attributi degli attori o degli eventi allo scopo di esplorare l'associazione tra i modelli relazionali e gli attributi. Allo stesso modo è possibile rappresentare, anche, la posizione dei gruppi di attori (o eventi) su piani fattoriali se sono disponibili le etichette dei cluster di appartenenza (71). Questi ultimi punti risultano di particolare interesse nel presente lavoro poichè costituiscono la base per la rappresentazione, in uno spazio metrico, delle posizioni derivanti da un'analisi di *two-mode* Blockmodeling.

Attori nello spazio degli eventi. Diversamente dall'applicazione dell'ACM ai dati numerici, in cui la rappresentazione delle righe come punti nello spazio delle variabili non è di particolare interesse, nell'applicazione a dati di affiliazione le righe corrispondono agli attori e, quindi, risulta di notevole interesse la loro rappresentazione in una mappa bidimensionale, costruita in base alle prime due colonne delle coordinate principale Φ (Eq. 3.4). In questa mappa, ogni attore corrisponde ad un punto e le distanze di tra i punti meglio approssima la distanza χ^2 tra i profili dell'attore nello spazio originario, e rappresentano le posizioni relative degli attori nel network. Due attori corrispondenti a due punti vicini nella mappa hanno simili *pattern* di partecipazione. Attori che corrispondono a punti vicino all'origine degli assi hanno *pattern* di partecipazione comuni (in quanto il profilo medio dell'attore è collocato vicino all'origine degli assi), mentre gli attori i cui punti corrispondenti sono collocati lontani dal centro hanno un modello di partecipazione inusuale. Inoltre, possono essere individuati anche attori isolati e gruppi di attori.

Eventi nello spazio degli attori Grazie all'approccio del *doubling*, ogni evento e_j nello spazio degli attori è rappresentato da due vettori opposti, corrispondenti ai due poli e_j^+ e e_j^- , giacenti sulla stessa direzione e passante per l'origine. La lunghezza di ogni segmento congiungente i due poli è proporzionale all'inverso del prodotto tra il tasso di partecipazione e di non-partecipazione (Eq 3.13). Le distanze $\delta_{e_j^+}$ e $\delta_{e_j^-}$ dei due poli dall'origine sono date da:

$$\delta_{e_j^+} = \left(1 - \frac{s_j}{n}\right); \quad \delta_{e_j^-} = \frac{s_j}{n} \quad (3.14)$$

Un modo alternativo di concepire tale approccio è quello di immaginare i punti finali di ogni segmento dell'evento come avente pesi proporzionali al tasso medio di partecipazione dei rispettivi poli. Se una distanza ($\delta_{e_j^+}$ o $\delta_{e_j^-}$) è molto più grande rispetto agli altri, l'evento è altamente polarizzata e quindi è raro o comune.

Inoltre, il coseno dell'angolo tra due segmenti di eventi esprime la correlazione tra i modelli di partecipazione negli eventi. Ne consegue che, se due segmenti degli eventi formano un piccolo angolo e presentano il polo positivo sullo stesso lato, gli eventi corrispondenti avranno modelli di partecipazione simili. Se i due

3.4 L'uso di punti supplementari per integrare informazioni esterne nella rappresentazione dell'ACM: Attributi e Posizioni

segmenti formano un angolo piccolo ma i poli positivi sono opposti, gli eventi avranno modelli di partecipazione opposti. Queste caratteristiche vengono conservate anche quando gli eventi vengono rappresentati sulle mappe bidimensionali ottenute con le prime due coordinate principali ψ (Eq. 3.5). Infine, nello spazio degli attori è possibile valutare la variabilità nella frequenza degli eventi. Per esempio, maggiore è la polarizzazione, minore è la variabilità nella partecipazione all'evento.

Rappresentazione congiunta degli attori e degli eventi. Gli attori e gli eventi possono essere rappresentati su una mappa bidimensionale congiunta attraverso il Biplot (108). In letteratura, per il caso dei dati numerici, sono state fatte varie proposte al fine di ottenere la rappresentazione più appropriata delle righe e delle colonne in un unico spazio. Il modo più semplice, come nell'AC semplice, può essere quello di sovrapporre la rappresentazione degli attori e degli eventi utilizzando, per entrambi, le prime due coordinate principali (Eq. 3.4 e 3.5) ottenendo una mappa simmetrica. Tuttavia, data l'asimmetria degli attori e degli eventi, D'Esposito e colleghi (71) propongono di mantenere tale asimmetria anche nella rappresentazione utilizzando il biplot asimmetrico in cui gli attori sono rappresentati in coordinate principali (Eq. 3.4) e gli eventi in coordinate standard (Eq. 3.7) (121). La direzione del vettore definita da ogni evento costituisce l'asse del biplot. Proiettando i punti che rappresentano gli attori su ciascun asse del biplot è possibile ottenere un valore approssimato del profilo di partecipazione all'evento e spiegare "perché" gli attori sono vicini, o lontani, nello spazio degli eventi.

3.4 L'uso di punti supplementari per integrare informazioni esterne nella rappresentazione dell'ACM: Attributi e Posizioni

Per caratterizzare gli attori e gli eventi attraverso i loro attributi, è possibile aggiungere, alla matrice delle variabili indicatrici \mathbf{Z} , un set di colonne (righe) corrispondenti alla codifica disgiuntiva completa delle variabili qualitative utilizzate come attributi (71). Allo stesso modo proponiamo di aggiungere nella matrice

3.4 L'uso di punti supplementari per integrare informazioni esterne nella rappresentazione dell'ACM: Attributi e Posizioni

delle variabili indicatrici i cluster degli attori e degli eventi (le posizioni per attori e dei cluster per eventi) derivanti dalla procedura di *Two-mode* Blockmodeling. I cluster vengono inseriti come set aggiuntivo di colonne (righe) corrispondenti alla codifica disgiuntiva completa delle variabili categoriali rappresentante il two-clustering. Nello specifico, gli attributi e i cluster degli eventi vengono aggiunti alla matrice di affiliazione come righe supplementari in cui vi è una variabile *dummy* per ogni modalità delle variabili categoriali degli eventi. Gli attributi e i cluster degli attori vengono aggiunti alla matrice di affiliazione come colonne supplementari in cui vi è una variabile *dummy* per ogni modalità delle variabili categoriali degli attori. In questo modo, partendo dalla matrice di affiliazione in codifica disgiuntiva completa \mathbf{Z} viene creata la matrice delle variabili indicatrici di analisi \mathbf{Z}^+ che assume la forma di una matrice a blocchi:

$$\mathbf{Z}^+ = \left(\begin{array}{c|c|c} \mathbf{Z} & \mathbf{A}^a & \mathbf{B}^a \\ \hline \mathbf{A}^e & \mathbf{0} & \mathbf{0} \\ \hline \mathbf{B}^e & \mathbf{0} & \mathbf{0} \end{array} \right), \quad (3.15)$$

dove

- \mathbf{Z} , è una matrice in codifica disgiuntiva completa che codifica la partecipazione, e la non partecipazione, di ogni attore ad ogni evento;
- \mathbf{A}^a e \mathbf{A}^e sono le matrici in codifica disgiuntiva completa degli attributi, rispettivamente, degli attori e degli eventi;
- \mathbf{B}^a e \mathbf{B}^e sono le matrici in codifica disgiuntiva completa delle partizioni risultanti dal Blockmodeling, ossia i cluster, rispettivamente, degli attori e degli eventi;

con

- $\mathbf{A}^a = (a_{ki}^a) = 1$, se l'attore i -esimo in riga presenta la caratteristica corrispondente alla k -sima modalità della variabile categoriale di attributo degli attori, 0, se non la possiede;
- $\mathbf{A}^e = (a_{kj}^e) = 1$, se l'evento j -esimo in colonna presenta la caratteristica corrispondente alla k -sima modalità della variabile categoriale di attributo degli eventi, 0, se non la possiede;

3.4 L'uso di punti supplementari per integrare informazioni esterne nella rappresentazione dell'ACM: Attributi e Posizioni

- $\mathbf{B}^a = (b_{ki}^a) = 1$, se l'attore i -esimo in riga appartiene al k -simo cluster di attori, o posizione, derivanti dal two-clustering stabilito empiricamente attraverso il *two-mode* Blockmodeling, 0, se non vi appartiene;
- $\mathbf{B}^e = (b_{kj}^e) = 1$, se l'evento j -esimo in colonna appartiene al k -simo cluster di eventi, derivanti dal two-clustering stabilito empiricamente, 0, se non vi appartiene;

Le informazioni aggiuntive non hanno alcuna influenza sulla soluzione dell'ACM, poiché non contribuiscono alla determinazione del δ_{ACM}^2 e degli assi principali (119). Più precisamente, le righe e le colonne supplementari possono essere viste come punti aggiuntivi con zero masse negli spazi dei profili delle righe o delle colonne.

Tuttavia, nonostante non contribuiscano alla determinazione analitica degli assi fattoriali, possono essere proiettati sul piano fattoriale come informazione aggiuntiva e le loro posizioni possono essere utilizzati per supportare l'interpretazione della configurazione. Inoltre, è possibile esplorare le relazioni tra i *pattern* relazionali emergenti dalla soluzione dell'ACM e i corrispondenti cluster stabiliti attraverso le procedure di Blockmodeling. In tal modo è possibile massimizzare il contributo informativo contenuto nei dati e arricchire l'interpretazione della soluzione finale, sia in un'ottica di ACM sia di Blockmodeling.

Infatti, il Blockmodeling contribuisce attraverso una classificazione, distinta, ma in un'unica procedura complessiva, degli attori e degli eventi fornendo una partizione, da un lato, delle posizioni occupate dagli attori e dei corrispondenti cluster degli eventi, sulla base dei *pattern* di relazioni che li caratterizzano (attori rispetto agli eventi a cui prendono parte e eventi rispetto agli attori che vi partecipano), e dall'altro, dei blocchi di relazioni sulla base dei *pattern* di connessioni esistenti tra le posizioni. Dal canto suo l'ACM, tra i vari contributi, offre al Blockmodeling la possibilità di rappresentare le posizioni degli attori, e dei cluster degli eventi, in uno spazio metrico, consentendo di analizzarne le associazioni con la struttura relazionale sottostante e valutarne la collocazione nello "spazio sociale" in termini di distanza. Inoltre, la presenza degli attributi descrittivi degli attori e degli eventi consente di ricostruire un "contesto informativo significativo" grazie

3.4 L'uso di punti supplementari per integrare informazioni esterne nella rappresentazione dell'ACM: Attributi e Posizioni

al quale le posizioni degli attori, e i cluster degli eventi, possono essere proficuamente caratterizzate ed interpretate. Anche se in modo, e su strutture di dati differenti, da quelle analizzate in questo lavoro, l'utilità di rappresentare in uno spazio metrico non solo le strutture relazionali di un network ma anche quelle derivanti dal Blockmodeling attraverso le tecniche dell'AC viene sottolineata anche in letteratura (44, 68, 110, 186, 192).

Nell'ACM della matrice delle variabili indicatrici (a blocchi \mathbf{Z}^+), le variabili supplementari possono essere rappresentate come medie dei punti dei rispondenti in coordinate principali (158). Ad esempio, attraverso la formula di transizione:

$$\mathbf{\Psi} = \mathbf{D}_e^{-1} \frac{\mathbf{Z}}{nm} \mathbf{\Phi} \mathbf{\Sigma}^{-1} = \mathbf{D}_e^{-1} \frac{\mathbf{Z}}{nm} \mathbf{\Gamma}, \quad (3.16)$$

che esprime le coordinate principali dell'evento come media ponderata delle coordinate standard dell'attore con pesi uguali ai profili dell'evento, le coordinate principali degli attributi dell'attore sono espressi come:

$${}_{as}\mathbf{\Psi} = {}_{as}\mathbf{D}_e^{-1} \frac{{}_{as}\mathbf{Z}}{nm} \mathbf{\Phi} \mathbf{\Sigma}^{-1} = {}_{as}\mathbf{D}_e^{-1} \frac{{}_{as}\mathbf{Z}}{nm} \mathbf{\Gamma}, \quad (3.17)$$

$${}_{bs}\mathbf{\Psi} = {}_{bs}\mathbf{D}_e^{-1} \frac{{}_{bs}\mathbf{Z}}{nm} \mathbf{\Phi} \mathbf{\Sigma}^{-1} = {}_{bs}\mathbf{D}_e^{-1} \frac{{}_{bs}\mathbf{Z}}{nm} \mathbf{\Gamma}, \quad (3.18)$$

dove

- ${}_{as}\mathbf{\Psi}$, ${}_{as}\mathbf{D}_e^{-1}$, ${}_{as}\mathbf{Z}$ sono, rispettivamente, le coordinate principali, la matrice dei pesi e la codifica disgiuntiva completa degli attributi degli attori che vengono aggiunti e,
- ${}_{bs}\mathbf{\Psi}$, ${}_{bs}\mathbf{D}_e^{-1}$, ${}_{bs}\mathbf{Z}$ sono, rispettivamente, le coordinate principali, la matrice dei pesi e la codifica disgiuntiva completa dei cluster degli attori ottenuti dal Blockmodeling.

L'utilizzo di informazioni aggiuntive può essere molto utile anche nell'analisi di grandi matrici affiliazione, in cui la valutazione del ruolo dei singoli attori o eventi è difficile o poco sensata. In questo caso, può essere più utile analizzare i *pattern* relazionali di sottogruppi di soggetti o eventi preservando tutte le informazione contenute nella matrice di affiliazione originaria (71).

3.5 Analisi delle Corrispondenze Multiple, Equivalenza e similarità strutturale: la metrica del Chi-quadrato

Nell'analisi delle reti di affiliazione, è possibile individuare un punto di notevole interesse per il presente lavoro, che costituisce la base per analizzare il legame tra l'ACM e Blockmodeling. Tale punto è costituito dalla possibilità offerta dall'ACM di valutare il grado di similarità dei profili degli attori/eventi nel network, che corrisponde alla valutazione dell'equivalenza strutturale tra gli attori/eventi (70, 71, 92, 95, 133, 142, 171). In un network *two-mode*, due attori sono strutturalmente equivalenti se partecipano esattamente gli stessi eventi e due eventi sono strutturalmente equivalenti se sono frequentati dagli stessi attori (166). Formalmente, dati due attori a_i e $a_{i'}$, e la proprietà di equivalenza strutturale \mathcal{S} si afferma che: $a_i \mathcal{S} a_{i'}$ se e solo se $f_{ij} = f_{i'j} \forall j$. Se due attori a_i e $a_{i'}$ sono strutturalmente equivalenti essi sono indistinguibili, e un attore equivalente può sostituire l'altro perché i due *pattern* relazionali sono gli stessi. Le stesse affermazioni valgono anche per gli eventi. Per misurare il grado di similarità tra attori/eventi in un network *two-mode* sono stati proposti due approcci principali: la prima basata sulla distanza Euclidea (56), l'altra sulla correlazione (45). Per un network *two-mode* descritto da una matrice di affiliazione \mathbf{F} , nell'approccio indiretto (*conversion approach*) le misure di distanza sopra illustrate sono applicate agli elementi delle matrici $\mathbf{F}'\mathbf{F}$ e $\mathbf{F}\mathbf{F}'$. Nell'approccio diretto, invece, in linea con Burt (56), le similarità tra gli attori/eventi possono essere misurate attraverso la distanza Euclidea tra i corrispondenti vettori numerici:

$$\begin{aligned}\delta_E^2(\mathbf{f}'_{a_i}, \mathbf{f}'_{a_{i'}}) &= \sum_{j=1}^m (f_{ij} - f_{i'j})^2, \\ \delta_E^2(\mathbf{f}_{e_j}, \mathbf{f}_{e_{j'}}) &= \sum_{i=1}^n (f_{ij} - f_{ij'})^2,\end{aligned}\tag{3.19}$$

dimostrano che le distanze nell'Eq. 3.19 sono linearmente legate alle distanze utilizzate sulle matrici $\mathbf{F}'\mathbf{F}$ e $\mathbf{F}\mathbf{F}'$ (71).

Se due attori/eventi sono strutturalmente equivalenti, le distanze nell'Eq. (3.19) sono uguali a zero, e viceversa. Le deviazioni da 0 possono essere uti-

3.5 Analisi delle Corrispondenze Multiple, Equivalenza e similarità strutturale: la metrica del Chi-quadrato

lizzate per interpretare un'equivalenza strutturale imperfetta. Un aumento della distanza tra attori/eventi corrisponde ad una riduzione della loro similarità.

Quando i metodi fattoriali come l'AC o l'ACM vengono utilizzati per rappresentare graficamente le similarità di attori o eventi, si ricorre alla distanza del χ^2 che misura gli scostamenti dall'equivalenza distributiva, invece che alle distanze Euclidee utilizzate per misurare il grado gli scostamenti dall'equivalenza strutturale in Eq. (3.19). È utile confrontare le differenze nelle rappresentazioni delle distanze del χ^2 e di quelle euclidee sui piani fattoriali. Nel caso dell'ACM la distanza del χ^2 è quella in Eq. (3.11) e (3.12). Un confronto delle due equazioni 3.19 con Eq. 3.11 e 3.12 mostra che, δ_{ACM}^2 , rispetto alla distanza Euclidea δ_E^2 , si basa su un sistema di pesi, che coinvolge le *size* (dimensioni) degli eventi. Tuttavia, la *size* dell'evento s_j ha come controparte il suo complemento $n - s_j$, e di conseguenza i pesi delle colonne sono in qualche modo bilanciati. La scelta di utilizzare l'ACM, piuttosto che l'AC, per analizzare e rappresentare graficamente la struttura di similarità di attori/eventi è legata ai risultati di recenti sviluppi (70, 71) che attraverso uno studio di simulazione vengono analizzate le differenze tra δ_e^2 , δ_{ACM}^2 e δ_{AC}^2 , considerando l'influenza della presenza di sotto-strutture (cluster), della dimensione e della densità del network su entrambe le misure e in confronto con quella Euclidea. Nel caso della distanza del χ^2 ottenuta tramite l'AC semplice, ossia quando l'algoritmo è applicato direttamente alla matrice \mathbf{F} considerata come una tabella di contingenza e f_{ij} sono le frequenze, la distanza del χ^2 , tra due profili di attori $\mathbf{p}_{a_i}, \mathbf{p}_{a_{i'}}$, è data da:

$$\delta_{AC}^2(\mathbf{p}_{a_i}, \mathbf{p}_{a_{i'}}) = (\mathbf{p}_{a_i} - \mathbf{p}_{a_{i'}})^T \mathbf{D}_e^{-1} (\mathbf{p}_{a_i} - \mathbf{p}_{a_{i'}}) = \sum_{j=1}^m \left(\frac{f_{ij}}{d_i} - \frac{f_{i'j}}{d_i'} \right)^2 \frac{L}{s_j}, \quad (3.20)$$

con $\mathbf{D}_e = \text{diag}(s_1/L, \dots, s_n/L)$, e L il numero totale dei legami, mentre, la distanza del χ^2 tra due profili di eventi, $\mathbf{p}_{e_j}, \mathbf{p}_{e_{j'}}$, è data da:

$$\delta_{AC}^2(\mathbf{p}_{e_j}, \mathbf{p}_{e_{j'}}) = (\mathbf{p}_{e_j} - \mathbf{p}_{e_{j'}})^T \mathbf{D}_a^{-1} (\mathbf{p}_{e_j} - \mathbf{p}_{e_{j'}}) = \sum_{i=1}^n \left(\frac{f_{ij}}{s_j} - \frac{f_{ij'}}{s_j'} \right)^2 \frac{L}{d_i}, \quad (3.21)$$

con $\mathbf{D}_a = \text{diag}(d_1/L, \dots, d_m/L)$ I risultati dello studio, in generale, mostrano che l'ACM è più simile dell'AC alla distanza δ_e^2 per tutti i casi analizzati, dato che l' R^2 presenta i valori più vicini ad uno e con una variabilità più bassa. Nel caso di

3.6 Blockmodeling, *Dual-projection Approach*, Analisi delle Corrispondenze Multiple

network sparsi, tale variabilità diventa minore per l'ACM rispetto all'AC. Questo comportamento viene associato alla grande variabilità nei le *degree* degli attori e alla le *size* degli eventi che, nelle configurazioni analizzate, produce una grande variabilità nei pesi dell'AC lasciando più stabili quelli dell'ACM. La metrica usata nell'ACM, ossia il sistema di pesi coinvolto nella distanza χ^2 , si rivela stabile rispetto ai *degree* degli attori e alle *size* degli eventi, e rispetto ad alcune caratteristiche del network come la densità e la presenza di cluster. Questo sistema di pesi, anche se potrebbe avere alcuni effetti indesiderati sulle mappe fattoriali, permette di preservare alcune preziose informazioni sulla struttura del network (70, 71).

3.6 Blockmodeling, *Dual-projection Approach*, Analisi delle Corrispondenze Multiple

In letteratura è possibile individuare un altro approccio allo studio delle strutture di similarità tra attori/eventi che porta a risultati simili a quelli visti sino ad ora per l'ACM. Si tratta del *Dual Projection Approach* sviluppato da Everett e Borgatti (91). Nella scelta tra un approccio diretto e uno indiretto per l'analisi delle *two-mode* network ad una approccio indiretto non necessariamente si accompagna una perdita di informazione a condizione che vengano analizzati entrambi i modi e che ne vengano integrati i risultati. Inoltre, anche nella definizione di una misura di (dis)similarità per l'analisi del "grado" di equivalenza strutturale tra i profili di attori/eventi possono essere individuate delle misure indirette che, basate su metodo appropriato di costruzione delle proiezioni, evitano il rischio di introdurre elementi di distorsione nella misura (91). Come si è detto, il grado di equivalenza strutturale viene definito come la misura in cui due nodi hanno legami e "non" legami con gli stessi altri (un terzo attore).

L'adozione del *Dual Projection Approach* per l'equivalenza strutturale, e quindi per il Blockmodeling, su dati *two-mode* implica la costruzione delle due proiezioni, il calcolo delle matrici di similarità per ciascuna di esse, una procedura di *Cluster Analysis* eseguita su ciascuna, e infine una mappatura delle partizioni risultanti sulla matrice di dati *two-mode* originari. In questa processo ciò che distingue tale approccio dagli altri approcci indiretti risiede, oltre che nel vincolo

di considerare entrambi i modi, nel metodo utilizzato per la costruzione delle proiezioni dei due modi. In quest'ottica, l'uso dei prodotti matriciali \mathbf{FF}' e $\mathbf{F}'\mathbf{F}$ per la costruzione delle proiezioni dei due modi rappresenta "solo una delle possibili classi di proiezioni" (91, p. 207), poichè si tratta della costruzione di matrici di similarità sulla base delle righe e delle colonne della matrice originale. Per l'analisi dell'equivalenza strutturale viene proposto una tecnica di proiezione che si ritiene rifletta in modo accurato la misura di equivalenza strutturale utilizzata. Piuttosto che utilizzare il coefficiente di Correlazione o la distanza Euclidea viene proposto l'utilizzo del *Proportion Matching* che viene considerato più appropriato data la natura binaria dei dati. Il *Proportion Matching* viene definito come il numero di volte in due profili hanno la stessa entrata (o 0 o 1) diviso per la lunghezza del profilo. Pertanto, due profili strutturalmente equivalenti assumono un valore di 1. In questo modo, per ottenere i gruppi strutturalmente equivalenti, diviene possibile effettuare il clustering direttamente sulle proiezioni (attraverso una procedura di *Combinatorial Optimization* che ricerca i cluster in modo da massimizzare il valore medio all'interno dei gruppi). Ciò consente di evitare la necessità di misurare la similarità dei profili sulle proiezioni e, successivamente, eseguire il clustering. Si ritiene che questo step aggiuntivo dei metodi indiretti tradizionali possa portare ad uno scostamento dati originari ed alla perdita informazioni strutturali (85).

Nel modo in cui viene definito il *Proportion Matching* corrisponde all'indice di similarità per dati binari di Sokal-Michener (187), anche noto come *Simple Matching Coefficient*. Tale indice implica una misura di similarità tra profili degli attori che attribuisce pari importanza al contributo informativo offerto dalle co-occorrenze e dalle co-assenze, in tal modo consente di tenere conto non solo delle co-partecipazioni ma anche delle "non" co-partecipazioni, definendo una misura di similarità più forte rispetto a quelle che considerano soltanto le co-partecipazioni. Nel contesto del presente lavoro è interessante notare che, data una matrice di affiliazione artificiale (Tab. 4.1) $\mathbf{F}(10 \times 10)$, con 10 attori e 10 eventi, i risultati ottenuti dal calcolo delle similarità tra i profili degli attori attraverso il *Proportion Matching* (Tab. 3.2) sono gli stessi di quelli ottenuti dal prodotto matriciale \mathbf{ZZ}' della una matrice di affiliazione \mathbf{F} in prospettiva *doubling*, ossia di \mathbf{Z} . Dove quest'ultima corrisponde alla matrice in codifica disgiuntiva completa ottenuta

3.6 Blockmodeling, *Dual-projection Approach*, Analisi delle Corrispondenze Multiple

	E1	E2	E3	E4	E5	E6	E7	E8	E9	E10
A1	1	1	1	1	0	1	0	0	0	0
A2	1	1	1	1	1	0	1	0	0	0
A3	1	1	1	1	0	1	0	0	0	0
A4	1	1	1	1	1	0	1	0	0	0
A5	0	0	0	0	1	1	1	0	1	0
A6	0	0	0	0	1	1	1	1	0	1
A7	0	0	0	0	1	1	1	0	1	0
A8	0	0	0	0	0	0	0	1	1	1
A9	0	0	0	0	0	0	0	1	1	1
A10	0	0	0	0	0	0	0	1	1	1

Tabella 3.1: Matrice *two-mode* \mathbf{F}

applicando l'ACM alla matrice di affiliazione di partenza considerandola come una matrice di variabili indicatrici (*Indicator Matrix*)(piuttosto che come una matrice di Burt). L'unica differenza risiede nel fatto che nel *Proportion Matching* la somma delle co-occorrenze e delle co-assenze vengono normalizzate attraverso le lunghezze dei profili, quelle che nel caso della matrice *doubled* corrispondono ai *degree* degli attori, che, in virtù del *doubling*, costituiscono una quantità costante pari a m , numero degli eventi.

Di conseguenza, normalizzando i prodotti matriciali di \mathbf{ZZ}' per il numero degli attori e degli eventi (mn) (ossia L , numero di egami) si ottengono gli stessi risultati del *Proportion Matching* (Tab. 3.3).

	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10
A1	0,0	0,7	1,0	0,7	0,3	0,2	0,3	0,2	0,2	0,2
A2	0,7	0,0	0,7	1,0	0,4	0,3	0,4	0,1	0,1	0,1
A3	1,0	0,7	0,0	0,7	0,3	0,2	0,3	0,2	0,2	0,2
A4	0,7	1,0	0,7	0,0	0,4	0,3	0,4	0,1	0,1	0,1
A5	0,3	0,4	0,3	0,4	0,0	0,7	1,0	0,5	0,5	0,5
A6	0,2	0,3	0,2	0,3	0,7	0,0	0,7	0,6	0,6	0,6
A7	0,3	0,4	0,3	0,4	1,0	0,7	0,0	0,5	0,5	0,5
A8	0,2	0,1	0,2	0,1	0,5	0,6	0,5	0,0	1,0	1,0
A9	0,2	0,1	0,2	0,1	0,5	0,6	0,5	1,0	0,0	1,0
A10	0,2	0,1	0,2	0,1	0,5	0,6	0,5	1,0	1,0	0,0

Tabella 3.2: Proiezione del set degli attori di \mathbf{F} tramite *Proportion Matching*

Tali risultati, inoltre, si differenziano da quelli ottenuti dal calcolo delle similarità tra i profili degli attori ottenuti dal prodotto matriciale \mathbf{FF}' della matrice di affiliazione non *doubled* (Tab. 3.4) mostrando una maggiore capacità di discriminazione tra i *pattern* di similarità dei profili degli attori.

3.6 Blockmodeling, *Dual-projection Approach*, Analisi delle Corrispondenze Multiple

	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10
A1	0	7	10	7	3	2	3	2	2	2
A2	7	0	7	10	4	3	4	1	1	1
A3	10	7	0	7	3	2	3	2	2	2
A4	7	10	7	0	4	3	4	1	1	1
A5	3	4	3	4	0	7	10	5	5	5
A6	2	3	2	3	7	0	7	6	6	6
A7	3	4	3	4	10	7	0	5	5	5
A8	2	1	2	1	5	6	5	0	10	10
A9	2	1	2	1	5	6	5	10	0	10
A10	2	1	2	1	5	6	5	10	10	0

Tabella 3.3: Matrice *one-mode* delle co-affiliazioni *doubled*: \mathbf{FF}'

	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10
A1	0	4	5	4	1	1	1	0	0	0
A2	4	0	4	6	2	2	2	0	0	0
A3	5	4	0	4	1	1	1	0	0	0
A4	4	6	4	0	2	2	2	0	0	0
A5	1	2	1	2	0	3	4	1	1	1
A6	1	2	1	2	3	0	3	2	2	2
A7	1	2	1	2	4	3	0	1	1	1
A8	0	0	0	0	1	2	1	0	3	3
A9	0	0	0	0	1	2	1	3	0	3
A10	0	0	0	0	1	2	1	3	3	0

Tabella 3.4: Matrice *one-mode* delle co-affiliazioni \mathbf{FF}'

Inoltre, sembra utile sottolineare che tali procedure, il *Proportion Matching* e il prodotto matriciale \mathbf{ZZ}' con fattore di normalizzazione pari a (mn) , $(\frac{\mathbf{ZZ}'}{nm})$, hanno una stretta connessione con la distanza Euclidea. In particolare, l'indice di Sokal-Michener, che ben riflette le caratteristiche che sottendono l'applicazione della codifica *doubling*, ha una relazione diretta con la distanza Euclidea, in quanto il complemento a 1 di quest'ultima ($1 - s_{SMij}$) costituisce la somma dei quadrati della distanza Euclidea (201), pertanto il complemento a 1 dell'indice di Sokal-Michener costituisce il quadrato della distanza Euclidea e $1 - s_{SMij} = d_{ij}^2$ quindi $\sqrt{1 - s_{SMij}} = d_{ij}$. Ai fini della presente trattazione risulta di notevole interesse il fatto che il calcolo delle distanze attraverso la metrica del χ^2 in ottica di ACM, δ_{ACM}^2 , produce dei risultati simili a quelle ottenute dall'indice Sokal-Michener, o il *Proportion Matching*, con l'aggiunta di un sistema di pesi più complesso per l'ACM che la rende stabile non solo rispetto ai *degree* degli attori ma anche alle *size* degli eventi. Tale peso è legato alla *size* dell'evento ed è pari al fattore $\frac{1}{s_j(1-s_j)}$

che costituisce l'inverso della varianza ($\frac{1}{p_j(1-p_j)}$) della variabile casuale Bernoulliana, $X \sim \text{Bern}(p)$, che, per ciascun evento e_j , ha probabilità (p_j) "partecipa" e probabilità ($1 - p_j$) "non partecipa" all'evento, con valore atteso $E(X) = p_j$ e varianza $\text{Var}(X) = p_j(1 - p_j)$. Questo fattore di correzione distingue anche δ_{ACM}^2 dalla distanza Euclidea (56) utilizzata come misura del grado di equivalenza strutturale nell'analisi di Blockmodeling, rendendo la prima misura più stabile rispetto alla *size* degli eventi (eventi a cui partecipano molti attori sono meno discriminanti rispetto ad eventi a cui partecipano poche attori). Nella distanza del χ^2 δ_{ACM}^2 , come per la distanza Euclidea δ_e^2 (54, 56), utilizzata nell'approccio Convenzionale al Blockmodeling, che operationalizza il grado di equivalenza (con la differenza del peso aggiuntivo) nell'equivalenza distribuzionale si ha una misura del grado in cui i due attori sono equivalenti (distributivamente equivalenti). Il concetto di equivalenza strutturale e distribuzionale risultano strettamente connessi, di conseguenza le distanze dei profili degli attori/eventi rappresentate sul piano fattoriale e i *pattern* di partecipazione derivanti sono molto vicini a quelle individuate attraverso un'analisi di Blockmodling diretto basato sul criterio di equivalenza strutturale, e non solo. Infatti, se si considera che il criterio di equivalenza regolare può essere visto come un rilassamento di quello strutturale e che i tipi di blocchi ideali alla base dell'equivalenza generalizzata risultano dal rilassamento dell'equivalenza regolare, non si esclude che alcuni di questi *pattern* di connessioni possano essere efficacemente individuati e rappresentati attraverso l'ACM.

I confronti sino ad ora esposti possono essere considerati come un ulteriore conferma dell'utilità della metrica coinvolta nell'ACM nell'analisi delle matrici *two-mode* e delle strutture relazionali sottostanti consentendo l'esplorazione della similarità nei *pattern* di partecipazione in un'ottica di equivalenza distribuzionale offrendo uno spazio multidimensionale appropriato per la valutazione delle distanze tra le posizioni e dei risultati del *two-mode* Blockmodeling. Inoltre in un'ottica di sviluppo di una strategia efficiente nella realizzazione di un'analisi di *two-mode* Blockmodeling, come si esplicherà meglio nei capitoli successivi, l'ACM consente di ottenere una serie di informazioni di supporto alle scelte da effettuare nel corso dell'analisi. Dal numero di cluster, alla individuazione di un modello di

3.6 Blockmodeling, *Dual-projection Approach*, Analisi delle Corrispondenze Multiple

struttura teorico per il blockmodel, alla scelta dei *constraint*, sino alla definizione di un "contesto informativo significativo" all'interno del quale rappresentare e caratterizzare i risultati del *two-mode* Blockmodeling.

Capitolo 4

Analisi delle Corrispondenze Multiple e Blockmodeling per Network *Two-mode*

4.1 Analisi delle Corrispondenze Multiple e Blockmodeling Esplorativo

Come si è visto nel capitolo secondo, nell'ambito del Blockmodeling generalizzato sono possibili due approcci, quello induttivo, o esplorativo, e quello deduttivo, o pre-specificato.

Nel caso del Blockmodeling esplorativo, l'analisi è basata solo sui dati relazionali e le scelte iniziali, riguardanti il numero di cluster per i due distinti set, degli attori e degli eventi, e il tipo di equivalenza sono soggettive. I tipi di equivalenza possibili per questo tipo di approccio sono l'equivalenza strutturale e quella regolare. I tipi di blocchi consentiti, rispetto al tipo di equivalenza selezionato (nullo e completo per l'equivalenza strutturale; nullo e regolare per l'equivalenza regolare) possono comparire ovunque nella struttura del blockmodel. Nell'analisi si procede per tentativi, valutando il numero di cluster e, eventualmente il tipo di equivalenza, più appropriato sulla base dei risultati della funzione criterio, che racchiude una misura della "bontà di adattamento" del blockmodel stabilito rispetto ai dati empirici, e in base alla significatività sostanziale della configurazione finale del blockmodel stabilito. Se non sono disponibili informazioni aggiuntive queste scelte iniziali sono arbitrarie e richiedono numerosi tentativi prima di giungere ad una soluzione ragionevole. Questa difficoltà aumenta con l'incremento del

numero di unità nel network e con la tendenza di quest'ultimi a diventare sparsi, ossia a presentare legami diradati. Questa proprietà del network è legata alla limitata capacità di un attore di mantenere legami con altri attori, o di prendere parte ad eventi, in quanto tale attività comporta un impiego di risorse, come il tempo, che non sono illimitate e possono comportare un costo (18). Un network "sparso" (*sparse network*) viene definito come un network in cui il numero di legami è dello stesso ordine di grandezza del numero di unità o il (*degree*) medio è molto più piccolo del numero di unità (12, 156). Questa tendenza diviene più accentuata man mano che la dimensione aumenta e si passa da network di medie dimensioni (fino ad 100 unità, *middle-size network*) a quelli di grandi (da qualche centinaia fino a migliaia unità, *large network*) (18). Nel caso dell'analisi di Blockmodeling Generalizzato, come si è visto, il problema della dimensione influisce anche sulla capacità della procedura di clustering di individuare tutte le partizioni ottimali possibili (83). Pertanto, in genere, tali tecniche tendono ad essere applicate su network di piccole dimensioni. Tuttavia, i vantaggi conoscitivi derivanti dall'analisi di Blockmodeling spingono alla ricerca di strategie e soluzioni che consentano la sua applicazione anche per network più grandi, complessi, e/o sparsi, in un'ottica sia di approcci diretti (tuttavia, non oltre qualche centinaia di unità) sia di approcci indiretti (quasi esclusivamente basati sull'equivalenza strutturale) (8, 98, 129, 170, 205). Inoltre, sono possibili, forme di pre-trattamento dei *large network* che prevedono la creazione di network ridotti, attraverso tecniche di decomposizione basate sulle misure di connettività (*weak components*, *strong components*, *symmetric-acyclic decomposition*, *bi-connected components*), i quali possono essere ulteriormente analizzati con le tecniche di Blockmodeling (98). Nel Blockmodeling esplorativo un altro problema emerge nella scelta della partizione finale in quanto sono possibili più partizioni modellate bene, coerenti con il criterio di equivalenza specificato e con un valore basso per la funzione criterio. In questo caso può essere complicato effettuare una scelta, soprattutto se le soluzioni sono differenti.

Nell'analisi dei network *two-mode* attraverso il *two-mode* Blockmodeling una delle principali direzioni in cui l'Analisi delle Corrispondenze Multiple può essere di supporto è nell'ambito dell'approccio esplorativo. In esso si propone di applicare il metodo di decomposizione di Eckart-Young (85) come metodo di riduzione ed

eseguire il Blockmodeling su una matrice di affiliazione di rango inferiore. Questa operazione consente di estrarre la struttura relazionale "forte", eliminando una parte delle inconsistenze, e facilita l'operazione di partizionamento. La scelta di tale strategia come fase esclusiva è da valutare con cautela, se non sconsigliabile, in quanto nella struttura "debole", che viene scartata insieme alle inconsistenze, potrebbero celarsi *pattern* relazionali di un certo interesse. Tale operazione di riduzione può essere suggerita come fase preliminare al Blockmodeling sul network originale e consente di ottenere una visione globale della struttura e di acquisire informazioni utili come guida per alcune scelte iniziali quali il numero di cluster. In questo lavoro l'applicazione del metodo verrà illustrata, in un primo momento su un network artificiale di piccole dimensioni, composta da 10 attori e 10 eventi, che, essendo di piccole dimensioni, consente di controllare e mostrare, passo dopo passo, gli esiti di tutti passaggi della procedura da noi proposta. In un secondo momento verranno utilizzati network reali di dimensioni più ampie (almeno 70 vertici). Quelli che verranno mostrati di seguito costituiscono i risultati relativi ai soli network analizzati. Pertanto, più che risultati definitivi relativi all'applicazione del metodo, quelli proposti, hanno lo scopo di evidenziare punti su cui le Tecniche dell'Analisi Fattoriale e quelle di Blockmodeling possono interagire, traendo beneficio dai rispettivi contributi. In quest'ottica, quindi, rappresentano dei suggerimenti alla base di ulteriori approfondimenti. ¹

4.1.1 Decomposizione di Eckart-Young e inconsistenze

Un primo modo per ottenere delle informazioni sulla struttura del network di affiliazione utile ad orientare le scelte per il *two-mode* Blockmodeling esplorativo (TMBE) è attraverso la ricostruzione della matrice originale operata mediante il metodo della Decomposizione in Valori Singolari (SVD), approssimando tale matrice ad una di rango inferiore che riproduca il più possibile la variabilità (inerzia) della matrice iniziale. Tale passaggio consente di estrarre la struttura forte del network, rimuovendone la struttura debole e il "rumore", ossia potenziali elementi

¹In questo lavoro non vengono utilizzate network di grandi dimensioni, tuttavia, si ritiene che il metodo possa essere utilmente applicato anche ad essi, si rimanda a futuri sviluppi la verifica di tale utilità.

di disturbo.

Il metodo per la Decomposizione in Valori Singolari (SVD) alla base dell'algoritmo dell'AC è definito per una matrice rettangolare \mathbf{X} ($n \times m$) di rango k , (in genere, con $k = \min\{n, m\}$), composta dalle matrici $(\mathbf{U}\mathbf{\Sigma}\mathbf{V}^T)$ dove $\mathbf{\Sigma}$ è la matrice diagonale dei valori singolari, e \mathbf{U} , \mathbf{V} sono, rispettivamente, quelle dei vettori singolari di sinistra e di destra. In breve, il metodo dell'SVD possiede un'importante proprietà in quanto permette di scomporre la matrice iniziale (nelle sue sub-matrici costitutive) e consente anche di ricostruirla a partire dai valori singolari e dai vettori singolari ottenuti (85). Infatti, è possibile ottenere un'approssimazione r -dimensionale di \mathbf{X} , $\hat{\mathbf{X}}_r$, attraverso la minimizzazione della somma dei residui al quadrato (116):

$$\min \|\mathbf{X} - \hat{\mathbf{X}}_r\|^2 \quad (4.1)$$

ottenuta attraverso

$$\hat{\mathbf{X}}_r = \mathbf{U}_r \mathbf{\Sigma}_r \mathbf{V}_m^T \quad (4.2)$$

oppure, in modo equivalente, attraverso

$$\hat{\mathbf{X}}_r = \sum_{i=1}^r (\sigma_i \mathbf{u}_i \mathbf{v}_m^T) \quad (4.3)$$

che esprime l'SVD come una somma di m matrici di rango 1. In essa si assume che i valori singolari siano ordinati dal più grande al più piccolo, cioè, $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_m > 0$. In tal modo il problema viene espresso come approssimazione dei minimi quadrati (117). La soluzione ottenuta in questo modo fornisce la migliore approssimazione di rango inferiore r (con $r < k$), di \mathbf{X} , in termini di minimi quadrati. I valori singolari sono ordinati in modo decrescente, pertanto, la somma dei primi due fornisce la migliore approssimazione di \mathbf{X} di rango $r = 2$ (116). Seguendo la procedura inversa è possibile ricostruire la matrice originale, sia interamente sia in parte. La ricostruzione implica la somma delle matrici della stessa dimensione dell'originale ma di rango 1 ottenute come prodotto di un vettore colonna $(n, 1)$ per un vettore di riga $(1, m)$ (136). Nel caso dell'ACM la matrice di affiliazione \mathbf{F} , quindi, può essere ricostruita in modo

esatto utilizzando le coordinate principali di riga e di colonna e i valori singolari su tutte le dimensioni contenuti nella matrice \mathbf{S} :

$$\mathbf{F} = \mathbf{D}_a^{1/2} \mathbf{U} \Sigma \mathbf{V}^T \mathbf{D}_a^{1/2} = (nm) \left[\frac{1}{n} \mathbf{1} \mathbf{1}^T \mathbf{D}_e + (\mathbf{D}_a^{1/2} (\mathbf{u} \mathbf{v}^T \sigma) \mathbf{D}_e^{1/2}) \right] \quad (4.4)$$

dove $(\mathbf{u} \mathbf{v}^T \sigma)$ rappresenta la matrice ricostruita sulla base di tutti i valori singolari e dei relativi vettori singolari di sinistra e di destra. Tuttavia, è possibile ricostruire una matrice \mathbf{F} di rango inferiore sulla base della somma delle matrici ricostruite a partire dai primi Q valori singolari, scelti sulla base dell'inerzia riprodotta. Pertanto, la matrice di rango inferiore sarà data da:

$$\mathbf{F}_{ric_Q} = (nm) \sum_{q=1}^Q \left[\frac{1}{n} \mathbf{1} \mathbf{1}^T \mathbf{D}_e + (\mathbf{D}_a^{1/2} (\mathbf{u}_q \mathbf{v}_q^T \sigma_q) \mathbf{D}_e^{1/2}) \right] \quad (4.5)$$

Nella procedura di ricostruzione come primo passo vengono rimosse una parte delle matrici, in genere quelle a cui sono associati i valori singolari, e quindi l'inerzia, più bassa. Questa operazione si basa sul presupposto che la maggior parte del contenuto informativo, e della variabilità, della matrice dei dati originali sia contenuta nelle prime matrici associate ai primi valori singolari. In termini di Blockmodeling, l'estrazione della matrice corrispondente ad un numero appropriato di valori singolari consente di ottenere una struttura più pulita di quella originale, ossia con un numero minore di inconsistenze. Questo può migliorare la capacità del Blockmodeling di individuare le partizioni, riducendo il numero di soluzioni alternative e il carico computazionale. Nel presente lavoro vengono mostrati, in un primo momento, i risultati relativi alla ricostruzione dei dati di una matrice di affiliazione artificiale che, essendo di piccole dimensioni, consente di controllare e mostrare, passo dopo passo, gli esiti di tutti i passaggi della procedura da noi proposta. In un secondo momento, verranno mostrati i risultati relativi ad una matrice reale. In entrambi i casi la matrice viene ricostruita a partire dai primi valori singolari, scelti in base alla percentuale di inerzia riprodotta. Il primo esempio si basa su una matrice di affiliazione $\mathbf{F}(10 \times 10)$, con 10 attori e 10 eventi. La matrice è stata strutturata seguendo uno schema segmentato in cui sono presenti tre gruppi di attori che sono legati a tre differenti gruppi di eventi.

4.1 Analisi delle Corrispondenze Multiple e Blockmodeling Esplorativo

Inoltre, sono stati aggiunti alcuni legami tra attori eventi esterni ai tre gruppi principali.

	E1	E2	E3	E4	E5	E6	E7	E8	E9	E10
A1	1	1	1	1	0	1	0	0	0	0
A2	1	1	1	1	1	0	1	0	0	0
A3	1	1	1	1	0	1	0	0	0	0
A4	1	1	1	1	1	0	1	0	0	0
A5	0	0	0	0	1	1	1	0	1	0
A6	0	0	0	0	1	1	1	1	0	1
A7	0	0	0	0	1	1	1	0	1	0
A8	0	0	0	0	0	0	0	1	1	1
A9	0	0	0	0	0	0	0	1	1	1
A10	0	0	0	0	0	0	0	1	1	1

Tabella 4.1: - Matrice *two-mode* \mathbf{F}

Per valutare la capacità della procedura di ricostruzione nella riproduzione della struttura del network, per ogni matrice ricostruita a partire dal primo valore singolare è stata condotta un'analisi attraverso il TMBE ed attraverso l'AC sulla matrice con codifica *doubling*¹. Nelle Figure 4.1 e 4.2 sono riportati i risultati del Blockmodeling e la sua rappresentazione grafica delle partizioni basate sul blockmodel attraverso un grafo bipartito.

Nello specifico, nella Figura 4.1, l'immagine (a) mostra la matrice permutata con le partizioni derivanti dal *two-mode* blockmodel, mentre l'immagine (b) mostra la matrice immagine del blockmodel, o anche detto blockmodel immagine, la matrice delle inconsistenze e il valore della funzione criterio. Il blockmodel è stato realizzato secondo il criterio dell'equivalenza strutturale, selezionando come numero iniziale tre cluster per gli attori e tre per gli eventi (*3, 3-partition*). Dato il criterio di equivalenza selezionato e il tipo di Blockmodeling, i tipi di blocchi ammessi sono quelli nulli e completi che possono comparire ovunque nella struttura del blockmodel. Dall'esame della matrice partizionata e dalla matrice immagine è possibile notare che sono stati correttamente individuati i tre cluster di attori legati a tre differenti cluster di eventi che definiscono tre blocchi completi. Nello specifico, le partizioni ottenute dal Blockmodeling sono $BA_{11}\{A5, A6, A7\}$ e $BE_{11}\{E5, E6, E7\}$ per il primo blocco; $BA_{22}\{A1, A2, A3, A4\}$ e $BE_{22}\{E1, E2,$

¹L'AC su una matrice con codifica *doubling*, nel caso di dati binari, è equivalente all'ACM su una matrice in codifica disgiuntiva completa.

4.1 Analisi delle Corrispondenze Multiple e Blockmodeling Esplorativo

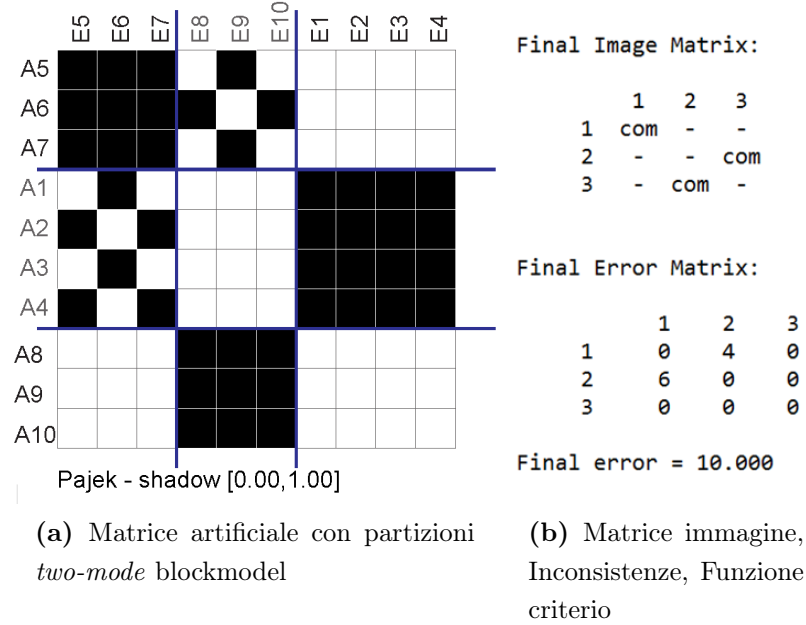


Figura 4.1: *Two-mode* Blockmodeling secondo l'equivalenza strutturale.

$E3, E4\}$ per il secondo blocco; $BA_{33}\{A8, A9, A10\}$ e $BE_{33}\{E8, E9, E10\}$ per il terzo blocco. Tale blockmodel, tuttavia, essendo basato sull'equivalenza strutturale, non riesce a modellare i blocchi definiti dai legami tra i cluster $BA_{12}\{A5, A6, A7\}$ e $BE_{21}\{E8, E9, E10\}$ e i cluster $BA_{21}\{A1, A2, A3, A4\}$ e $BE_{12}\{E5, E6, E7\}$. La matrice immagine che sintetizza la struttura del network analizzato, in cui ogni entrata corrisponde al valore assunto dal blocco corrispondente, mostra chiaramente i tre blocchi completi. La funzione criterio restituisce 10 inconsistenze, che, come visibile nella matrice delle inconsistenze corrispondono ai 6 legami nel blocco dato da BA_{12} e BE_{21} e 4 in quelli dati da BA_{21} e BE_{12} .

Prima di procedere con i risultati ottenuti dalla procedura di ricostruzione sono necessarie alcune osservazioni sulla struttura di questo blockmodel. Il tipo di inconsistenze rivelate nel blockmodel non costituiscono dei veri e propri "errori", intesi come legami caratterizzati dall'assenza di un chiaro *pattern* di connessioni (167), ma, piuttosto, il segnale di un inappropriato criterio di equivalenza alla base del blockmodel, o meglio non appropriato a descrivere il modello di connessioni dei blocchi definiti da BA_{12} e BE_{21} e da BA_{21} e BE_{12} . In questo caso vengono definite come inconsistenze indotte dal modello (*model-induced inconsistencies*)

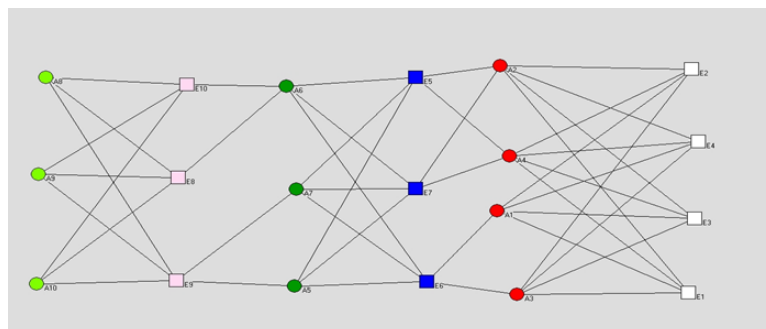


Figura 4.2: Grafo bipartito con partizioni del *two-mode* blockmodel matrice artificiale. Equivalenza strutturale.

(167). Ricordando la definizione dell'equivalenza regolare per la quale due attori sono regolarmente equivalenti se sono legati ad attori equivalenti, ossia dello stesso tipo. Inoltre, considerando i tipi di blocchi ammessi per tale criterio, ossia nullo e regolare (almeno un legame per riga e per colonna) risulta facile riconoscere nei blocchi definiti da BA_{12} e BE_{21} e da BA_{21} e BE_{12} un modello di connessioni coerente con l'equivalenza regolare. Tuttavia, per ragioni di semplicità nell'analisi esemplificativa condotta, è stata scelta una procedura di Blockmodeling diretto esplorativo elementare in cui è possibile solo la scelta del criterio di equivalenza e del numero di gruppi. Tra i possibili criteri di equivalenza è stata scelta quella strutturale poiché si è rivelata l'unica in grado di fornire un'unica partizione coerente con i dati empirici. Infatti, nelle analisi preliminari l'adozione dell'equivalenza regolare come criterio per la costruzione del blockmodel ha condotto a molteplici partizioni ben modellate, differenti tra loro e poco interpretabili¹. Tali risultati hanno suggerito l'utilizzo di un criterio di equivalenza più restrittivo quale quello dell'equivalenza strutturale².

¹Questi rappresentano alcuni dei problemi legati all'utilizzo del tipo di blocco *regular* che tende ad essere compatibile con numerose configurazioni di legami, pertanto, viene suggerito un utilizzo cauto (83)).

²Una struttura come quella della matrice di affiliazione in esame può essere modellata in modo efficace nell'ambito del Blockmodeling Generalizzato, che, consentendo di combinare più tipi di blocchi ideali, corrispondenti a diversi criteri di equivalenza, permette di ricercare i tipi di blocchi ideali che meglio rispecchiano tali modelli di connessioni

4.1.2 Blockmodeling su matrice di affiliazione di rango inferiore

La procedura di decomposizione e ricostruzione della matrice di affiliazione artificiale originale $\mathbf{F}(n \times m)$ parte (i) dall'applicazione dell'SVD alla matrice $\mathbf{Z}(n \times 2m)$, ossia la matrice \mathbf{F} in codifica *doubled*¹. Il secondo passo (ii) prevede la ricostruzione di una matrice di rango inferiore \mathbf{F}_{ric_Q} a partire dalla matrice \mathbf{S} . Nei passi successivi la matrice ricostruita \mathbf{F}_{ric_Q} viene analizzata attraverso il TM-BE basato sull'equivalenza strutturale (iii) e attraverso l'AC (iv). Il punto (iii) viene eseguito per analizzare l'effetto della ricostruzione sull'individuazione del blockmodel e il (iv) per una rappresentazione analitica e visuale della struttura ricostruita. Entrambe nel tentativo di comprendere come i due diversi metodi ricostruiscono la struttura sottostante del network. Ogni operazione viene ripetuta per la matrice ricostruita a partire, rispettivamente, dal primo valore singolare \mathbf{F}_{ric_1} ; dalla somma del primo e del secondo valore singolare \mathbf{F}_{ric_2} ; dalla somma del primo, del secondo e del terzo valore singolare \mathbf{F}_{ric_3} . Nello specifico, per la matrice di affiliazione artificiale, adottata come struttura esemplificativa, le matrici ricostruite attraverso la procedura riproducono, rispettivamente, il 58%, l'82% e il 93%² dell'inerzia totale, calcolata come rapporto tra inerzia riprodotta dalla matrice ricostruita sull'inerzia totale della matrice originale:

$$J_Q = \frac{\sum_{\alpha \leq Q} \sigma^2}{\sum_{\alpha=m} \sigma^2} \quad (4.6)$$

Ricostruzione di \mathbf{F}_{ric_1} in base al primo valore singolare. La Tabella 4.2 riporta i risultati della matrice ricostruita a partire dal primo valore singolare che riproduce il 58% dell'inerzia totale. Nella matrice vengono riportati in corsivo i legami assenti nella struttura originale erroneamente ricostruiti e in grassetto i legami presenti nella struttura originale ma non ricostruiti dalla procedura. I legami all'interno del primo gruppo di attori ed eventi (A1, A2, A3, A4 e E1, E2, E3, E4) e del terzo (A8, A9, A10 e E8, E9, E10) vengono ricostruiti correttamente.

¹Si ricorda che, nel caso di dati binari, la matrice *doubled* è equivalente alla matrice delle variabili indicatrici in codifica disgiuntiva completa.

²Il valore riportato corrisponde al tasso di inerzia non rivalutato.

4.1 Analisi delle Corrispondenze Multiple e Blockmodeling Esplorativo

Non vengono ricostruiti i legami del gruppo centrale (A5, A6, A7 e E5, E6, E7). Mentre, per i legami esterni ai due gruppi, oltre a quelli esistenti, ne vengono ricostruiti alcuni non presenti nella struttura originale. Infatti, il primo gruppo di attori presenta una partecipazione per tutti gli eventi del secondo gruppo di eventi: A1, A2, A3, A4 sono legati ad E5, E6, E7 piuttosto che A1 e A3 con E6 ed A2 e A4 con E5 e E7 effettivamente presenti nella struttura originale. Lo stesso accade per il secondo gruppo per i quali vengono ricostruiti i legami di partecipazione per tutti gli eventi del terzo gruppo di eventi (A5, A6, A7 sono legati ad E8, E9, E10 piuttosto che A5 e A7 con E9 ed A6 con E8 e E10). È possibile che tale ricostruzione sia dovuta al particolare *pattern* di partecipazione degli attori A5, A6, A7 che presentano legami con eventi a cui sono legati sia il primo gruppo di attori sia il secondo rendendolo più difficile da individuare attraverso un solo valore singolare.

Dai risultati dell'analisi di TMBE riportati nella Figura 4.3 è visibile una partizione in due blocchi completi definiti dai cluster degli attori e degli eventi (A5, A6, A7, A8, A9, A10 con E8, E9, E10 e A1, A2, A3, A4 con E1, E2, E3, E4, E5, E6, E7). Il Blockmodeling non restituisce inconsistenze.

Ricostruzione di F_{ric_2} in base al primo e secondo valore singolare. La Tabella 4.3 riporta i risultati della matrice ricostruita a partire dal primo e dal secondo valore singolare che riproduce l'82% dell'inerzia totale. Come nella matrice precedente, vengono riportati in corsivo i legami assenti nella struttura originale

	E1N	E1Y	E2N	E2Y	E3N	E3Y	E4N	E4Y	E5N	E5Y	E6N	E6Y	E7N	E7Y	E8N	E8Y	E9N	E9Y	E10N	E10Y
A1	0	1	0	1	0	1	0	1	<i>0</i>	<i>1</i>	0	1	<i>0</i>	<i>1</i>	1	0	1	0	1	0
A2	0	1	0	1	0	1	0	1	0	1	<i>0</i>	<i>1</i>	0	1	1	0	1	0	1	0
A3	0	1	0	1	0	1	0	1	<i>0</i>	<i>1</i>	0	1	<i>0</i>	<i>1</i>	1	0	1	0	1	0
A4	0	1	0	1	0	1	0	1	0	1	<i>0</i>	<i>1</i>	0	1	1	0	1	0	1	0
A5	1	0	1	0	1	0	1	0	1	0	1	0	1	0	<i>0</i>	<i>1</i>	0	1	<i>0</i>	<i>1</i>
A6	1	0	1	0	1	0	1	0	1	0	1	0	1	0	0	1	<i>0</i>	<i>1</i>	0	1
A7	1	0	1	0	1	0	1	0	1	0	1	0	1	0	<i>0</i>	<i>1</i>	0	1	<i>0</i>	<i>1</i>
A8	1	0	1	0	1	0	1	0	1	0	1	0	1	0	0	1	0	1	0	1
A9	1	0	1	0	1	0	1	0	1	0	1	0	1	0	0	1	0	1	0	1
A10	1	0	1	0	1	0	1	0	1	0	1	0	1	0	0	1	0	1	0	1

Tabella 4.2: Matrice di affiliazione ricostruita, 1 valore singolare, F_{ric_1} . Inerzia: 0,58. Corsivo: legami assenti ricostruiti; grassetto: legami presenti non ricostruiti.

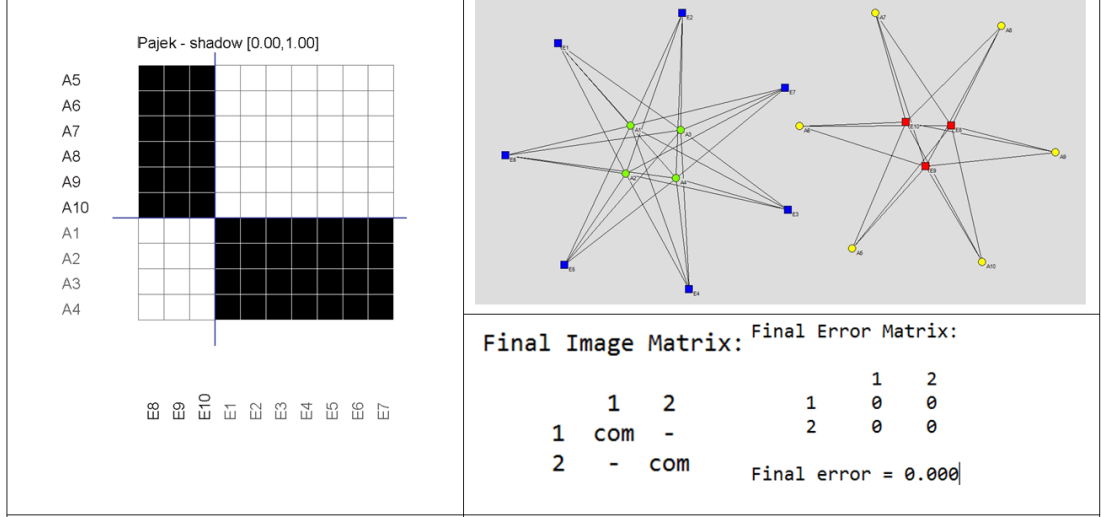


Figura 4.3: *Two-mode* Blockmodeling sulla matrice di affiliazione ricostruita, 1 valore singolare, \mathbf{F}_{ric1} . Inerzia: 0,58. Matrice con partizioni blockmodel. Grafo bipartito con partizioni. Matrice immagine, Matrice inconsistenze, Funzione criterio.

erroneamente ricostruiti e in grassetto i legami presenti nella struttura originale ma non ricostruiti dalla procedura. Tutti i legami relativi ai tre gruppi di attori ed eventi (A1, A2, A3, A4 con E1, E2, E3, E4; A5, A6, A7 con E5, E6, E7; A8, A9, A10 con E8, E9, E10) vengono ricostruiti correttamente. Non vengono ricostruiti correttamente i legami del gruppo A1, A2, A3, A4 con l'evento E6, in particolare è presente il legame di A1 e A3 con E6 mancante nella struttura originale e manca, invece, quello tra A2 e A4 con E6. Allo stesso modo, per l'attore A6 compare un legame non esistente con l'evento E9 ma sono assenti quelli con gli eventi E8 e E10. Dall'aggiunta del contributo informativo della matrice associata al secondo valore singolare è possibile scorgere con una certa chiarezza la struttura del network rilevando un ruolo critico dell'attore A6 e dell'evento E6.

I risultati dell'analisi di TMBE riportati nella Figura 4.4 mostrano l'individuazione di una partizione in tre blocchi completi definiti dai cluster degli attori e quelli degli eventi principali (A1, A2, A3, A4 con E1, E2, E3, E4; A5, A6, A7 con E5, E6, E7; A8, A9, A10 con E8, E9, E10). Il blockmodel presenta un valore per la funzione criterio pari a 9 (corrispondenti alle inconsistenze generate

4.1 Analisi delle Corrispondenze Multiple e Blockmodeling Esplorativo

	E1N	E1Y	E2N	E2Y	E3N	E3Y	E4N	E4Y	E5N	E5Y	E6N	E6Y	E7N	E7Y	E8N	E8Y	E9N	E9Y	E10N	E10Y
A1	0	1	0	1	0	1	0	1	1	0	1	0	1	0	1	0	1	0	1	0
A2	0	1	0	1	0	1	0	1	0	1	0	1	0	1	1	0	1	0	1	0
A3	0	1	0	1	0	1	0	1	1	0	1	0	1	0	1	0	1	0	1	0
A4	0	1	0	1	0	1	0	1	0	1	0	1	0	1	1	0	1	0	1	0
A5	1	0	1	0	1	0	1	0	0	1	0	1	0	1	1	0	0	1	1	0
A6	1	0	1	0	1	0	1	0	0	1	0	1	0	1	1	0	0	1	1	0
A7	1	0	1	0	1	0	1	0	0	1	0	1	0	1	1	0	0	1	1	0
A8	1	0	1	0	1	0	1	0	1	0	1	0	1	0	0	1	0	1	0	1
A9	1	0	1	0	1	0	1	0	1	0	1	0	1	0	0	1	0	1	0	1
A10	1	0	1	0	1	0	1	0	1	0	1	0	1	0	0	1	0	1	0	1

Tabella 4.3: Matrice di affiliazione ricostruita, 1 e 2 valore singolare, \mathbf{F}_{ric_2} . Inerzia: 0,82. Corsivo: legami assenti ricostruiti; grassetto: legami presenti non ricostruiti.

dai legami nei blocchi teoricamente nulli) e due soluzioni. Queste ultime sono sostanzialmente simili, con l'unica differenza data dall'evento E9 assegnato o al cluster di eventi (E8, E9, E10), come in Figura 4.4, o al cluster (E5, E6, E7). Confrontando le due matrici immagine e i due blockmodel costruiti a partire dal primo e dalla somma del primo e del secondo valore singolare, si può osservare che la struttura ricostruita dai primi due valori singolari è uguale alla struttura del blockmodel sulla matrice originaria. I tre blocchi completi, inoltre, sono definiti da cluster di attori ed eventi che sono gli stessi del two-clustering di partenza. I legami ricostruiti male confluiscono nella funzione criterio ma non alterano la struttura della partizione del blockmodel e del clustering.

Dai risultati dell'analisi di AC (in prospettiva *doubling*) (ACd)¹, riportati nella Tabella 4.4 è visibile una suddivisione molto simile a quella individuata dal TMBE. Osservando i contributi assoluti e relativi degli attori e degli eventi nella determinazione delle dimensioni nel nuovo sottospazio di riferimento e dall'analisi delle mappe fattoriali (Figura 4.5) relative alle prime due dimensioni, è possibile individuare due modelli di partecipazione e di non partecipazione in opposizione. Sulla prima dimensione sono visibili, da un lato, gli attori A2 e A4 in opposizione ad A8, A9, A10, e dall'altro, sulla seconda dimensione, A1 e A3 in opposizione ad A5, A6 e A7. A differenza del Blockmodeling l'ACd risente maggiormente dei

¹Si ricorda, ancora una volta, che l'AC su una matrice con codifica *doubling*, nel caso di dati binari, è equivalente all'ACM su una matrice in codifica disgiuntiva completa, pertanto, i risultati delle due tecniche sono equivalenti.

4.1 Analisi delle Corrispondenze Multiple e Blockmodeling Esplorativo

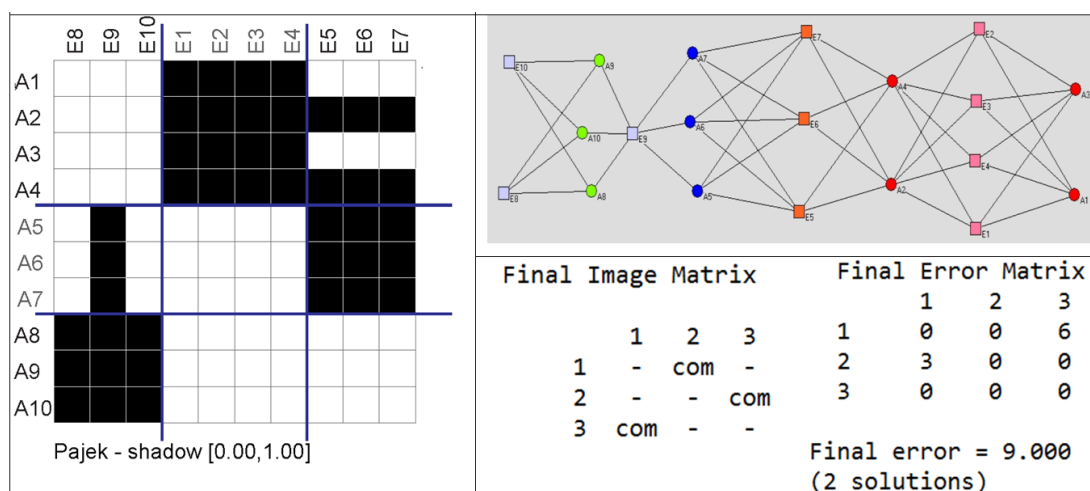


Figura 4.4: *Two-mode* Blockmodeling sulla matrice di affiliazione ricostruita, 1 e 2 valore singolare, F_{ric2} . Inerzia: 0,82. Matrice con partizioni blockmodel. Grafo bipartito con partizioni. Matrice immagine, Matrice inconsistenze, Funzione criterio.

legami ricostruiti male accentuano le differenze tra attori A1 e A3 ed A2 e A4 che risultano rappresentate su due dimensioni differenti. Tuttavia, questo significa anche che l'ACd, rispetto al Blockmodeling basato sull'equivalenza strutturale, è più sensibile all'individuazione di unità che sono presenti su più di una dimensione e che, quindi, non possono essere assegnate in modo preciso ad un unico gruppo.

Ricostruzione di F_{ric3} in base ai primi tre valori singolari. La Tabella 4.5 riporta i risultati della matrice ricostruita a partire ai primi tre valori singolari che riproduce il 93% dell'inerzia totale. In questa matrice, tutti i legami relativi ai tre gruppi di attori ed eventi (A1, A2, A3, A4 con E1, E2, E3, E4; A5, A6, A7 con E5, E6, E7; A8, A9, A10 con E8, E9, E10) vengono ricostruiti correttamente e vengono ricostruiti correttamente anche i legami del gruppo A1, A2, A3, A4 con l'evento E6. L'unico per il quale si rileva una ricostruzione parziale è l'attore A6 che presenta un legame mancante con E6 e uno non esistente con l'evento E9. Anche in questa matrice permane il ruolo critico dell'attore A6.

I risultati dell'analisi di TMBE riportati nella Figura 4.6 mostrano un cambiamento nell'individuazione della struttura dovuto all'attore A6. L'assegnazione del

4.1 Analisi delle Corrispondenze Multiple e Blockmodeling Esplorativo

	Dim1	Dim2	Dim3		Dim1	Dim2	Dim3
<i>Eigenv.</i>	<i>0,59</i>	<i>0,36</i>	<i>0,03</i>	<i>Eigenv.</i>	<i>0,59</i>	<i>0,36</i>	<i>0,03</i>
A1	8,97	14,67	16,35	E1Y	8,59	2,34	1,07
A2	17,18	0,81	22,01	E2Y	8,59	2,34	1,07
A3	8,97	14,67	16,35	E3Y	8,59	2,34	1,07
A4	17,18	0,81	22,01	E4Y	8,59	2,34	1,07
A5	0,75	17,85	4,72	E5Y	1,09	11,41	4,17
A6	0,75	17,85	4,72	E6Y	1,09	11,41	4,17
A7	0,75	17,85	4,72	E7Y	1,09	11,41	4,17
A8	15,14	5,16	3,03	E8Y	7,66	4,21	23,13
A9	15,14	5,16	3,03	E9Y	5,73	1,56	0,71
A10	15,14	5,16	3,03	E10Y	7,66	4,21	23,13

(a) Contributi attori

(b) Contributi eventi

	Dim1	Dim2	Dim3		Dim1	Dim2	Dim3
<i>Eigenv.</i>	<i>0,59</i>	<i>0,36</i>	<i>0,03</i>	<i>Eigenv.</i>	<i>0,59</i>	<i>0,36</i>	<i>0,03</i>
A1	0,46	0,47	0,05	E1Y	0,85	0,14	0,01
A2	0,89	0,02	0,07	E2Y	0,85	0,14	0,01
A3	0,46	0,47	0,05	E3Y	0,85	0,14	0,01
A4	0,89	0,02	0,07	E4Y	0,85	0,14	0,01
A5	0,06	0,91	0,02	E5Y	0,13	0,84	0,03
A6	0,06	0,91	0,02	E6Y	0,13	0,84	0,03
A7	0,06	0,91	0,02	E7Y	0,13	0,84	0,03
A8	0,81	0,17	0,01	E8Y	0,65	0,22	0,13
A9	0,81	0,17	0,01	E9Y	0,85	0,14	0,01
A10	0,81	0,17	0,01	E10Y	0,65	0,22	0,13

(c) Coseni quadrati attori

(d) Coseni quadrati eventi

Tabella 4.4: AC con *doubling* sulla matrice di affiliazione ricostruita, 1 e 2 valore singolare, \mathbf{F}_{ric_2} . Inerzia: 0,82.

legame con E9 e l'assenza di quello con E6, genera un'alterazione della configurazione dei legami e della partizione individuata dal TMBE. Infatti, pur restando invariato il numero di blocchi completi individuati, cambia la composizione di alcuni dei cluster che li definiscono. In particolare, cambia quella del primo cluster degli attori associati agli eventi E5, E6, E7, infatti A6 viene ricollocato nel cluster degli attori A8, A9, A10 legati al cluster degli eventi E8, E9, E10. Questa alterazione suggerisce che l'aggiunta di ogni nuovo valore singolare comporta l'aggiunta di nuove informazioni ma anche di potenziali elementi di disturbo.

Dai risultati dell'analisi di ACd, riportati nella Tabella 4.6 è possibile rilevare una suddivisione molto simile a quella precedente, che, in parte riproduce quella individuata dal TMBE e, in parte, mette in luce alcune caratteristiche che nel

4.1 Analisi delle Corrispondenze Multiple e Blockmodeling Esplorativo

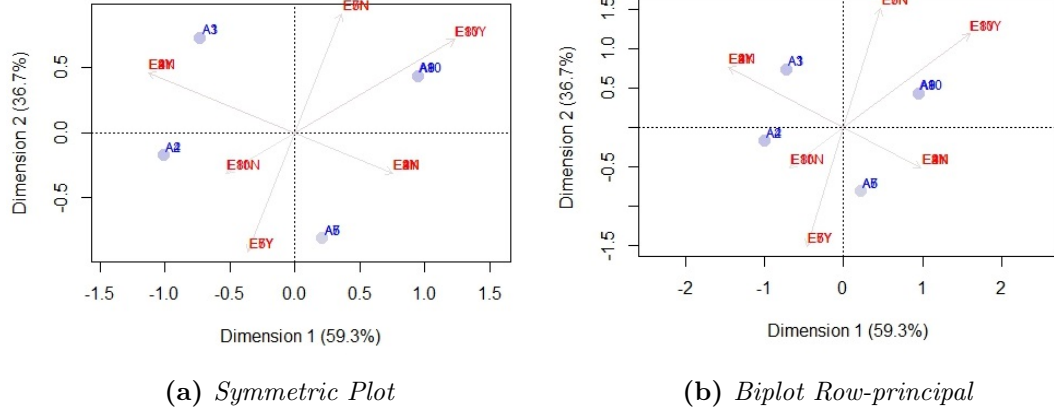


Figura 4.5: AC con *doubling* sulla matrice di affiliazione ricostruita, 1 e 2 valore singolare, \mathbf{F}_{ric_2} . Inerzia: 0,82.

TMBE vanno perse. Osservando i contributi assoluti e relativi degli attori e degli eventi nella determinazione delle dimensioni nel nuovo sottospazio di riferimento e dall'analisi delle mappe fattoriali (Figura 4.7) relative alle prime due dimensioni, è possibile fare una serie di osservazioni. In primo luogo, sono ravvisabili nuovamente i due modelli distinti di partecipazione e di non partecipazione ma in un'opposizione meno marcata. Sulla prima dimensione sono visibili i due modelli di partecipazione associati agli attori A1, A2, A3, A4 in opposizione ad A8, A9, A10. Su questa dimensione viene riprodotto anche il modello di connessioni

	E1N	E1Y	E2N	E2Y	E3N	E3Y	E4N	E4Y	E5N	E5Y	E6N	E6Y	E7N	E7Y	E8N	E8Y	E9N	E9Y	E10N	E10Y
A1	0	1	0	1	0	1	0	1	1	0	0	1	1	0	1	0	1	0	1	0
A2	0	1	0	1	0	1	0	1	0	1	1	0	0	1	1	0	1	0	1	0
A3	0	1	0	1	0	1	0	1	1	0	0	1	1	0	1	0	1	0	1	0
A4	0	1	0	1	0	1	0	1	0	1	1	0	0	1	1	0	1	0	1	0
A5	1	0	1	0	1	0	1	0	0	1	0	1	0	1	1	0	0	1	1	0
A6	1	0	1	0	1	0	1	0	0	1	1	0	0	1	0	1	<i>0</i>	<i>1</i>	0	1
A7	1	0	1	0	1	0	1	0	0	1	0	1	0	1	1	0	0	1	1	0
A8	1	0	1	0	1	0	1	0	1	0	1	0	1	0	0	1	0	1	0	1
A9	1	0	1	0	1	0	1	0	1	0	1	0	1	0	0	1	0	1	0	1
A10	1	0	1	0	1	0	1	0	1	0	1	0	1	0	0	1	0	1	0	1

Tabella 4.5: Matrice di affiliazione ricostruita, 1, 2 e 3 valore singolare, \mathbf{F}_{ric_3} . Inerzia: 0,93. Corsivo: legami assenti ricostruiti; grassetto: legami presenti non ricostruiti.

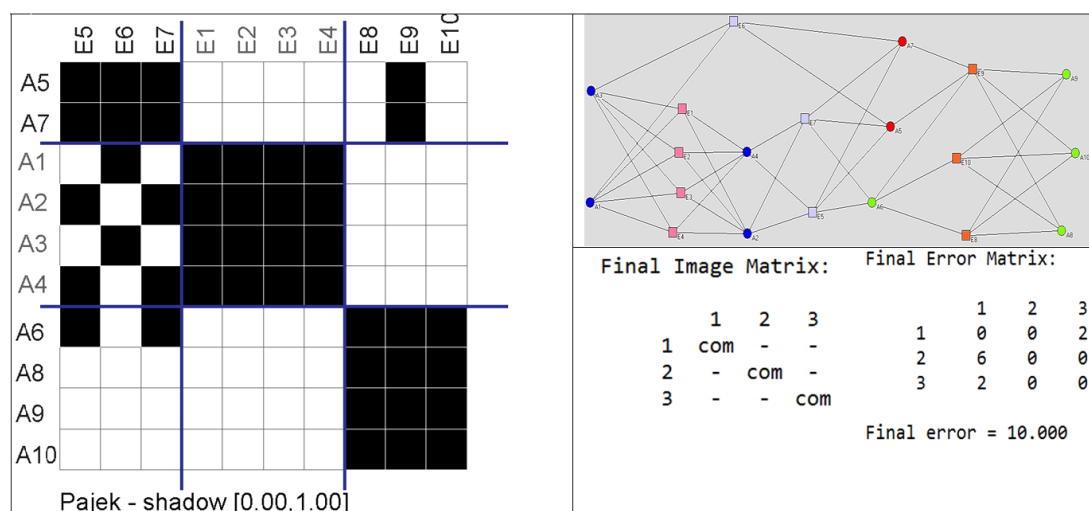


Figura 4.6: *Two-mode* Blockmodeling sulla matrice di affiliazione ricostruita, 1, 2 e 3 valore singolare, \mathbf{F}_{ric3} . Inerzia: 0,93. Matrice con partizioni blockmodel. Grafo bipartito con partizioni. Matrice immagine, Matrice inconsistenze, Funzione criterio

dell'attore A6 che a causa della ricostruzione parziale dei legami con E6 ed E9 si differenzia da A5 e A7. Sulla seconda dimensione, invece, viene riprodotto il modello di partecipazione di A5 e A7 che risulta distinto dai precedenti.

A questo punto è necessario notare che la differenziazione tra A1 e A3 ed A2 e A4 è più attenuata, essendo stati correttamente ricostruiti anche i legami esterni al blocco principale di appartenenza. Pertanto, i quattro attori risultano correttamente riprodotti sulla stessa dimensione. È interessante notare che l'informazione sui legami esterni, più forte per A2 e A4, non viene persa ma riprodotta sulla seconda dimensione. Questa constatazione conferma la sensibilità dell'ACd a trattenere e restituire in modo dettagliato le caratteristiche della struttura relazionale originaria, ricostruendo le strutture principali sulle prime dimensioni e le restanti sulle dimensioni successive.

Ricostruzione della matrice di affiliazione delle coproduzioni teatrali in base ai primi quattro valori singolari. Come secondo esempio di analisi di TMBE su una matrice di rango inferiore si riporta la matrice di affiliazione

4.1 Analisi delle Corrispondenze Multiple e Blockmodeling Esplorativo

	Dim1	Dim2	Dim3		Dim1	Dim2	Dim3
<i>Eigenv.</i>	<i>0,62</i>	<i>0,23</i>	<i>0,13</i>	<i>Eigenv.</i>	<i>0,62</i>	<i>0,23</i>	<i>0,13</i>
A1	14,04	8,83	10,88	E1Y	8,89	1,35	1,01
A2	13,67	1,40	19,79	E2Y	8,89	1,35	1,01
A3	14,04	8,83	10,88	E3Y	8,89	1,35	1,01
A4	13,67	1,40	19,79	E4Y	8,89	1,35	1,01
A5	0,80	25,51	12,13	E5Y	0,21	17,70	4,92
A6	9,16	4,13	14,13	E6Y	1,30	1,82	34,91
A7	0,80	25,51	12,13	E7Y	0,21	17,70	4,92
A8	11,27	8,13	0,09	E8Y	6,88	4,48	4,13
A9	11,27	8,13	0,09	E9Y	5,92	0,90	0,67
A10	11,27	8,13	0,09	E10Y	6,88	4,48	4,13

(a) Contributi attori

(b) Contributi eventi

	Dim1	Dim2	Dim3		Dim1	Dim2	Dim3
<i>Eigenv.</i>	<i>0,62</i>	<i>0,23</i>	<i>0,13</i>	<i>Eigenv.</i>	<i>0,62</i>	<i>0,23</i>	<i>0,13</i>
A1	0,71	0,17	0,12	E1Y	0,92	0,05	0,02
A2	0,74	0,03	0,23	E2Y	0,92	0,05	0,02
A3	0,71	0,17	0,12	E3Y	0,92	0,05	0,02
A4	0,74	0,03	0,23	E4Y	0,92	0,05	0,02
A5	0,06	0,74	0,20	E5Y	0,03	0,84	0,13
A6	0,63	0,11	0,21	E6Y	0,14	0,07	0,77
A7	0,06	0,74	0,20	E7Y	0,03	0,84	0,13
A8	0,78	0,21	0,00	E8Y	0,71	0,18	0,09
A9	0,78	0,21	0,00	E9Y	0,92	0,05	0,02
A10	0,78	0,21	0,00	E10Y	0,71	0,18	0,09

(c) Coseni quadrati attori

(d) Coseni quadrati eventi

Tabella 4.6: AC con *doubling* sulla Matrice di affiliazione ricostruita, 1, 2 e 3 valore singolare, \mathbf{F}_{ric_3} . Inerzia: 0,93.

relativa alle coproduzioni teatrali nel Sistema Teatrale Campano, nella stagione 2011-2012. Tale struttura di dati, insieme a quelle relative ad altre tre stagioni, sarà analizzata in modo più dettagliato nel corso dell'ultimo capitolo. La matrice delle coproduzioni teatrali è una matrice di affiliazione con 33 organizzazioni (attori) e 37 spettacoli in coproduzione (eventi). Nel network sono presenti 84 legami, una densità dello 0,07, con un tasso medio di partecipazione degli attori (*affiliation average rate*) pari a 2,55, e un numero medio di attori per evento (*event average sizes*) pari a 2,27. Nella Figura 4.8 si riportano i risultati del TMBE (6-6)-*partition*, con 6 gruppi di attori e 6 di eventi, ottenuti attraverso il criterio dell'equivalenza strutturale, sulla matrice di affiliazione originale. Come si può notare dalle immagini (a) e (b) la procedura partiziona il network in 4

4.1 Analisi delle Corrispondenze Multiple e Blockmodeling Esplorativo

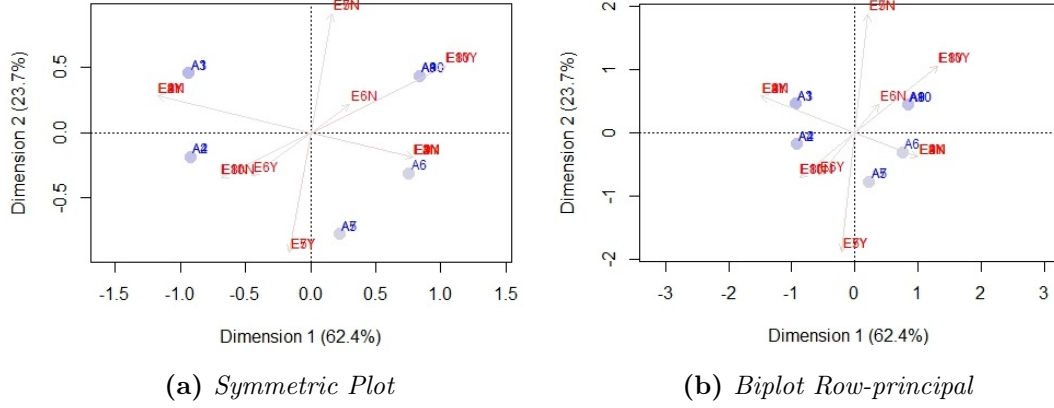


Figura 4.7: AC con *doubling* sulla Matrice di affiliazione ricostruita, 1, 2 e 3 valore singolare, \mathbf{F}_{ric_3} . Inerzia: 0,93.

blocchi completi con un valore per la funzione criterio pari a 44 inconsistenze. Il primo blocco è definito dal primo cluster di attori (Bracco e Totò) associati al primo cluster di eventi. Il secondo blocco è definito dal secondo cluster di attori (Campania Festival e Stabile Napoli) associati al quarto cluster di eventi. Il terzo blocco è definito dal terzo cluster di attori (Teatro Azione e Magazzini) associati al quinto cluster di eventi. Infine, il quarto blocco è definito dal quinto cluster di attori (Nuvole) associati al terzo cluster di eventi. Al di là della composizione dei cluster, dall'analisi della matrice partizionata nell'immagine (a), è evidente che il criterio dell'equivalenza strutturale risulta insufficiente per modellare correttamente la struttura relazionale. In effetti, ad esempio, le compagnie ProgMuseo e Nuvole mostrano un *pattern* di partecipazione comune, allo steso modo le compagnie Babylon e Fool. Tuttavia, anche con un numero diverso di cluster iniziali, tali blocchi non vengono identificati. Come si può notare i network caratterizzati da legami radi, definiti network sparsi, pongono una serie di problemi all'applicazione delle tecniche di Blockmodeling diretto, soprattutto, a quello basato sull'equivalenza strutturale. Nel network in esame, infatti, il grado di densità risulta molto basso (0,07) e questo impedisce l'individuazione ottimale dei blocchi sulla base di questo criterio di equivalenza che, per la definizione di un blocco, richiede che tutti componenti appartenenti ad un cluster di attori siano legati a

4.1 Analisi delle Corrispondenze Multiple e Blockmodeling Esplorativo

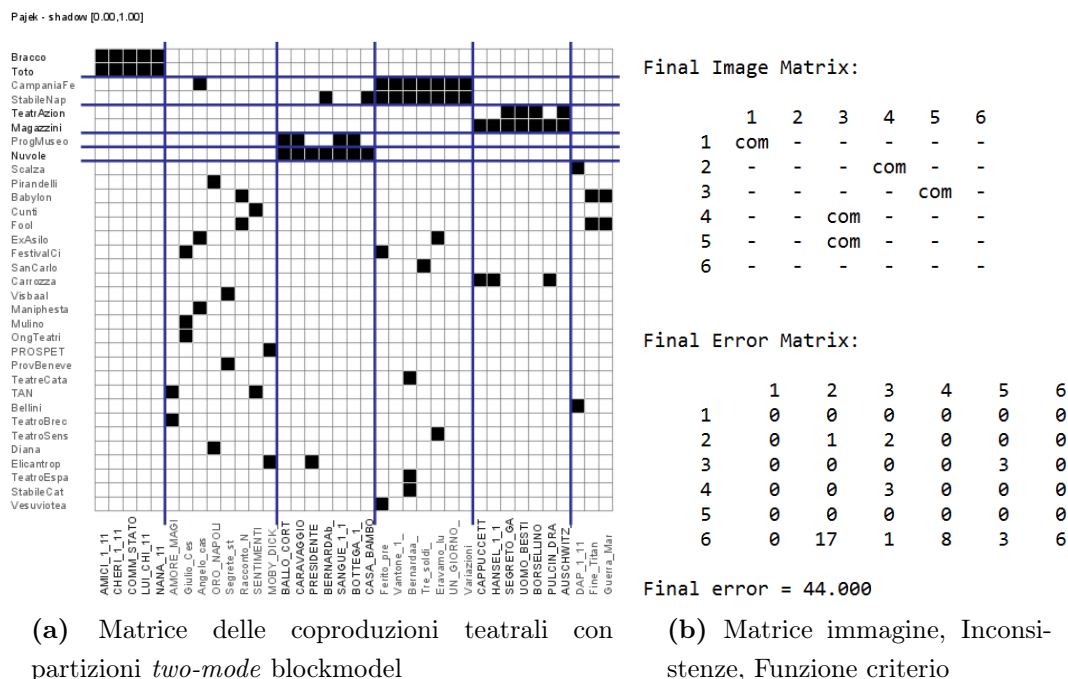


Figura 4.8: *Two-mode* Blockmodeling secondo l'equivalenza strutturale.

tutti i componenti di un cluster degli eventi. Nel caso di network sparsi, quindi, l'operazione di corretta identificazione dei modelli di connessione diviene più complicata. A tal proposito una possibile soluzione è il ricorso al Blockmodeling basato sull'equivalenza generalizzata (83) che consente di indicare, per ogni blocco, uno o più tipi di blocchi ideali al fine di individuare quelli che maggiormente riflettano i modelli di connessione reali. Tuttavia, quando la dimensione della rete aumenta, è necessario eseguire diversi tentativi di analisi prima di giungere all'individuazione del numero di cluster appropriato e del modello di struttura. In tal senso, un passo preliminare al Blockmodeling generalizzato può essere dato dall'analisi di una matrice di rango inferiore che consenta, prima di tutto, di estrarre la struttura 'forte' del network, tale da chiarire quali possano essere il numero e la forma delle sotto-strutture che lo compongono.

Prima di procedere alla presentazione del blockmodel stabilito a partire dalla matrice ridotta, sembra utile illustrare, passo dopo passo, i risultati della ricostruzione della matrice originaria attraverso i primi autovalori, in modo da evi-

denziarne le caratteristiche.

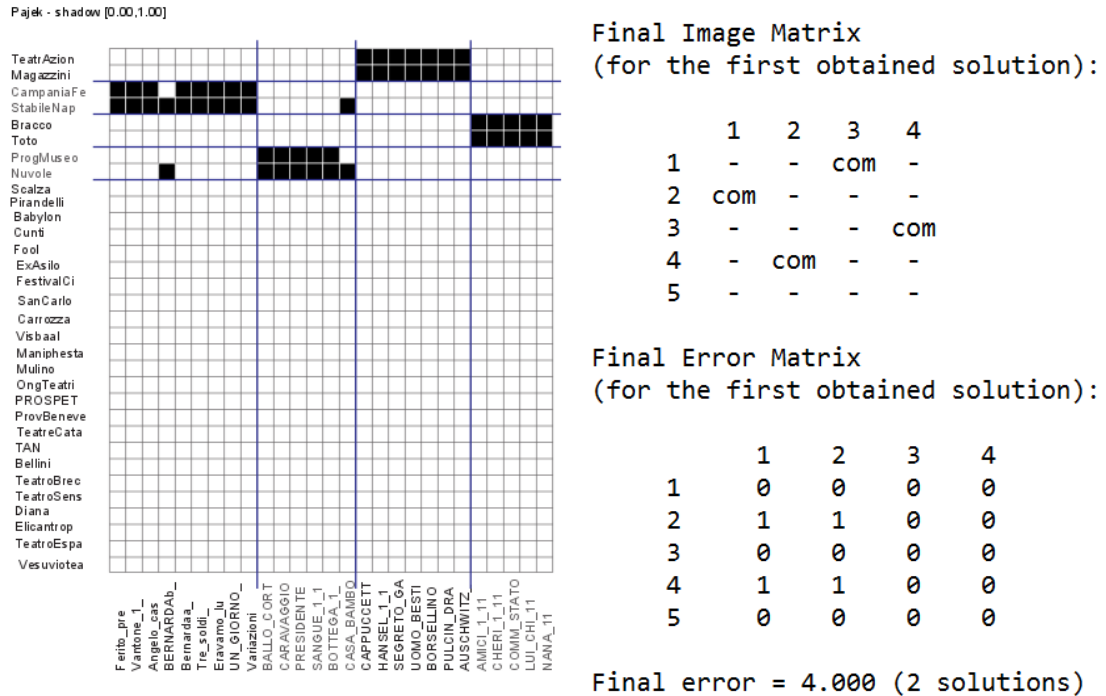
Nell'analisi delle singole matrici ricostruite è stato riscontrato che ognuna di esse, come nella matrice artificiale, ricostruisce una parte della struttura della matrice originaria e, nelle prime di esse, si concentra la struttura 'forte' del network. In particolare, in ognuna delle matrici iniziali viene ricostruito il blocco di attori ed eventi più importanti. Nell'ordine, nella prima matrice, ricostruita a partire dal primo valore singolare, è stato ricostruito il blocco più forte dato dal cluster formato da Campania Festival e Stabile Napoli e dal cluster delle coproduzioni corrispondenti. Nella seconda matrice, ricostruita a partire dai primi due valori singolari, è stato ricostruito il blocco costituito dal cluster formato da Teatro azione e Magazzini e dal cluster delle coproduzioni corrispondenti. Nella terza matrice, ricostruita a partire dai primi tre valori singolari, è stato ricostruito il blocco definito dal cluster dato da Totò e Bracco e dal cluster delle coproduzioni corrispondenti. Nella quarta matrice, ricostruita a partire dai primi quattro valori singolari, è stato ricostruito il blocco costituito dal cluster composto da Nuvole e ProgMuseo e dal cluster delle coproduzioni corrispondenti. Nella quinta matrice, ricostruita a partire dai primi cinque valori singolari, è stato ricostruito il blocco definito dal cluster composto da Babylon e Fool e dal cluster delle coproduzioni corrispondenti. Nella sesta matrice, ricostruita a partire dai primi sei valori singolari, al cluster costituito da Teatro azione e Magazzini si aggiunge anche Carrozza e scompaio tutti gli errori di ricostruzione. In effetti, com'è stato notato nell'applicazione del metodo di ricostruzione alla matrice artificiale, ogni volta che si aggiunge un nuovo valore singolare si ha una sorta di riconfigurazione dei legami. Nello specifico è stato notato che vengono ricostruiti per primi i blocchi che, in termini di blockmodel, si possono definire completi, ossia quelli in cui tutti i componenti di un cluster di attori sono associati a tutti le componenti di un cluster di coproduzioni. Nel caso in esame, questo avviene quando le compagnie di un cluster condividono tutte le coproduzioni di un certo cluster e pertanto possono essere definite come strutturalmente equivalenti. In un secondo momento, il processo di ricostruzione tende ad attribuire agli attori inclusi in uno stesso cluster gli stessi legami di coproduzione, sia internamente sia esternamente al blocco principale. Pertanto, emerge una tendenza a generare blocchi compatti,

4.1 Analisi delle Corrispondenze Multiple e Blockmodeling Esplorativo

completi con alcuni legami erroneamente ricostruiti o mancanti che, via via, andando avanti con la ricostruzione vengono riassegnati correttamente. Per quanto riguarda i legami di coproduzione sparsi, che non hanno un *pattern* forte, essi vengono ricostruiti per ultimi. Nella ricostruzione della matrice delle coproduzioni teatrali, si è scelto di fermarsi ai primi 4 valori singolari, sufficienti a riprodurre il 57% dell'inerzia totale della matrice originale. Nella Figura 4.9 viene riportato il risultato del TMBE sulla matrice ricostruita, dopo la SVD, attraverso i primi 4 valori singolari che, insieme, riproducono il 57% (non rivalutata) dell'inerzia totale della matrice di affiliazione originale. Com'è visibile dalle immagini (a) e (b) la procedura di Blockmodeling individua le partizioni del blockmodel, assegnando correttamente tutti i componenti ai cluster appropriati, senza dividere Campania Festival e Stabile Napoli in due cluster distinti, come nel caso precedente. Inoltre, il blockmodel è stato estratto in un unico tentativo, portando a due soluzioni sostanzialmente identiche, con l'unica differenza nell'assegnazione dello spettacolo Bernarda.b o al primo cluster di coproduzioni, associate al secondo cluster di attori, o al secondo cluster di coproduzioni, associato al quarto cluster di attori. Dalle immagini emerge una chiara struttura segmentata, composta da quattro blocchi completi. Tuttavia è necessario notare l'esistenza di alcuni legami di coproduzione assegnati ma non esistenti nella matrice originale. Si tratta della coproduzione Angelo casa con Stabile Napoli; Bernarda.b assegnato a Campania Festival; Presidente legato a ProgMuseo; infine, Cappuccetto, Hansel e Pulcin-drage che vengono legati a Teatro Azione.

Gli esiti della ricostruzione possono essere osservati anche dai risultati dell'ACM eseguita sulla matrice delle coproduzioni ridotta. Dalla Tabella 4.7 emerge chiaramente la suddivisione della struttura nei quattro gruppi, che concorrono, ognuno, a caratterizzare una dimensione differente della soluzione fattoriale. In particolare, sul primo asse si collocano Stabile Napoli e Campania Festival; sul secondo, Teatro azione e Magazzini; sul terzo, Nuvole e ProgMuseo; infine, sul quarto, si collocano Totò e Bracco. Per ognuno di essi, sulla stessa dimensione, si collocano anche le specifiche coproduzioni in cui sono coinvolti. La possibilità di sintetizzare la struttura fondamentale della matrice originale può generare dei vantaggi in termini di analisi di Blockmodeling. Infatti, questa operazione conduce ad una riduzione delle inconsistenze che facilita l'individuazione delle

4.1 Analisi delle Corrispondenze Multiple e Blockmodeling Esplorativo



(a) Matrice delle coproduzioni teatrali ricostruita con partizioni *two-mode* blockmodel

(b) Matrice immagine, Inconsistenze, Funzione criterio

Figura 4.9: *Two-mode* Blockmodeling sulla matrice di affiliazione ricostruita attraverso i primi 4 valori singolari. Inerzia: 0,57. Matrice con partizioni blockmodel. Matrice immagine, Matrice inconsistenze, Funzione criterio

partizioni, con un vantaggio sia in termini computazionali sia nel trattamento di network ampie e/o sparse. La scelta di tale strategia come fase esclusiva è da valutare con cautela, se non sconsigliabile, in quanto nella struttura "debole", che viene scartata insieme alle inconsistenze, potrebbero celarsi *pattern* relazionali interessanti. Tale operazione di riduzione può essere condotta come fase preliminare esplorativa prima del Blockmodeling sul network originale. Questa operazione consente di ottenere una visione globale della struttura e di acquisire informazioni utili come guida per alcune scelte iniziali sia in un'ottica di blockmodel esplorativo sia, e soprattutto, come si vedrà in seguito, di Blockmodeling pre-specificato. Essa, infatti può essere condotta come step preliminare esplorati-

4.2 Analisi delle Corrispondenze Multiple e Blockmodeling Pre-specificato

eta2	Dim1	Dim2	Dim3	Dim4
	0,32	0,26	0,20	0,18
Ferito.preludio	0,96	CAPPUCETTO 0,98	BALLO.CORTE.1 0,93	AMICI.1 0,59
Vantone.1	0,96	HANSEL.1 0,98	CARAVAGGIO.1 0,93	CHERI.1 0,59
Angelo.casa	0,96	SEGRETO.GAIA.1 0,98	PRESIDENTE 0,93	COMM.STATO 0,59
Bernardaa	0,96	UOMO.BESTIA.1 0,98	SANGUE.1 0,93	LUI.CHI 0,59
Tre.soldi	0,96	BORSELLINO.1 0,98	BOTTEGA.1 0,93	NANA 0,59
Eravamo.lupi	0,96	PULCIN.DRAGO 0,98	BERNARDAb 0,24	
UN.GIORNO.1	0,96	AUSCHWITZ 0,98	CASA.BAMBOLA.1 0,24	
Variazioni.Mito	0,96			
BERNARDAb	0,40			
CASA.BAMBOLA.1	0,40			
cos2	Dim1	Dim2	Dim3	Dim4
	0,32	0,26	0,20	0,18
CampaniaFestival	0,88	TeatrAzione 0,98	ProgMuseo 0,85	Bracco 0,59
StabileNap	0,96	Magazzini 0,98	Nuvole 0,88	Totò 0,59

Tabella 4.7: ACM sulla matrice delle coproduzioni teatrali 2011-2012 ricostruita attraverso i primi 4 valori singolari. Inerzia: 0,57. Coproduzioni e organizzazione rilevanti per dimensione.

vo per definire il numero di cluster, un modello ideale di struttura del blockmodel e i vincoli.

4.2 Analisi delle Corrispondenze Multiple e Blockmodeling Pre-specificato

Accanto al Blockmodeling esplorativo, un secondo approccio è quello del Blockmodeling pre-specificato che può essere ottenuto attraverso la specificazione, parziale o totale, di alcuni elementi del clustering e del blockmodel prima dell'analisi (11). La specificazione può essere realizzata sulla base delle conoscenze aggiuntive possedute dal ricercatore. Tali conoscenze possono riguardare gli attori (ad esempio, attributi demografici e ruoli) oppure la struttura dei legami tra le posizioni, ossia, ipotesi sulla struttura radicate nelle conoscenze teoriche (struttura gerarchica, *core-periphery*, ecc.)(83) . L'idea centrale, quindi, è legata al fatto che il Blockmodeling possa essere utilizzato non solo in modo induttivo (esplorativo) ma anche deduttivo (confermativo), integrando e valorizzando le conoscenze disponibili sull'oggetto di studio.

Come si è visto nel secondo capitolo (2.4.3) la pre-specificazione può assumere varie forme e gradi di rigidità e seconda degli elementi che vengono specificati e dei vincoli che vengono imposti alla struttura del clustering e del blockmodel (ad es., composizione, parziale o completa, dei cluster, collocazione, parziale o completa,

4.2 Analisi delle Corrispondenze Multiple e Blockmodeling Pre-specificato

dei tipi di blocchi ideali in specifici punti della struttura del blockmodel, ecc.). In questo contesto, l'ACM può essere di supporto al *two-mode* Blockmodeling Pre-specificato in quanto, insieme al numero di cluster, può fornire ulteriori informazioni da utilizzare nella specificazione di un modello di struttura del blockmodel che riproduca appropriatamente i modelli di connessione presenti nei dati. Come si è detto la pre-specificazione del modello richiede delle informazioni aggiuntive, a quelle relazionali, di carattere teorico o empirico. L'ACM, consentendo l'introdurre di informazioni aggiuntive nella procedura di analisi, risulta molto efficace a tale scopo. Quest'ultima fornisce un utile supporto alla ricostruzione del contesto della struttura relazionale e alla definizione dei possibili vincoli (*constraint*) da inserire nel modello, con conseguenze positive sulla stabilità delle soluzioni finali.

4.2.1 Ricostruzione del contesto, selezione del numero di cluster, vincoli

Per apprezzare le opportunità offerte dall'analisi esplorativa preliminare attraverso l'ACM, verranno illustrati i risultati ottenuti sulla matrice originale delle coproduzioni teatrali nella stagione 2011-2012.¹ Nella matrice delle coproduzioni teatrali sono state aggiunte numerose informazioni relative alle caratteristiche degli attori, ossia le compagnie, e degli evento e gli spettacoli in coproduzione. Gli attributi descrittivi degli attori riguardano una serie di caratteristiche quali: il tipo di organizzazione; il tipo di città (capoluogo di regione, capoluogo di provincia, e l'area provinciale); la dimensione e il tipo di sede del teatro principale (teatro storico, teatro moderno, costruzione non-teatrale riadattata per ospitare spettacoli dal vivo). Per quanto riguarda gli attributi descrittivi delle coproduzioni sono stati considerati: il genere teatrale (commedia, *cabaret*, contemporaneo, *musical*, teatro giovanile, sperimentale, ecc.); le caratteristiche dello script² (originale, rivisitato, esistente non rivisitato); la presenza di elementi di multidisciplinarietà (combinazione di forme di arte, media e comunicazione); la presenza, o meno, di interazione con gli spettatori. Tutti questi attributi sono

¹Tale matrice di affiliazione è stata utilizzata nel corso del presente capitolo per illustrare la procedura di decomposizione tramite la SVD e ricostruzione tramite il metodo di Eckart-Young.

²Lo script si riferisce l'opera letteraria o dramma su cui si basa una certa produzione

4.2 Analisi delle Corrispondenze Multiple e Blockmodeling Pre-specificato

Dim1			Dim2			Dim3		
17,07	coord.	cos2	14,79	coord.	cos2	13,76	coord.	cos2
CampaniaFestival	1,20	0,60	TeatrAzione	0,75	0,33	ProgMuseo	-1,06	0,65
StabileNap	1,47	0,73	Magazzini	1,26	0,53	Nuvole	-1,43	0,68
			Bracco	-1,08	0,54			
			Toto	-1,08	0,54			
Attributi attori						Attributi attori		
T.Prod.h	0,75	0,15	SC.Medium	1,54	0,15	T.Edu.C.instit	-2,86	0,26
T.Theat.Fest	1,51	0,15	SC.Large	-1,00	0,18			
Coproduzioni			Coproduzioni			Coproduzioni		
Vantone.1	3,22	0,67	AMICI.1	-2,81	0,51	BALLO.CORTE.1	-3,36	0,73
UN.GIORNO.1	3,22	0,67	CHERI.1	-2,81	0,51	CARAVAGGIO.1	-3,36	0,73
Variazioni.Mito	3,22	0,67	COMM.STATO	-2,81	0,51	SANGUE.1	-3,36	0,73
Tre.soldi	2,22	0,49	LUI.CHI	-2,81	0,51	BOTTEGA.1	-3,36	0,73
Eravamo.lupi	1,73	0,41	NANA	-2,81	0,51	PRESIDENTE	-2,22	0,32
BERNARDAb	2,48	0,40	SEGRETO.GAIA.1	2,62	0,44			
CASA.BAMBOLA.1	2,48	0,40	UOMO.BESTIA.1	2,62	0,44			
Ferito.preludio	1,67	0,39	BORSELLINO.1	2,62	0,44			
Bernardaa	1,34	0,32	AUSCHWITZ	2,62	0,44			
			CAPPUCETTO	2,26	0,33			
			HANSEL.1	2,26	0,33			
			PULCIN.DRAGO	2,26	0,33			
Attributi coproduz.						Attributi coproduz.		
G.CONT	0,78	0,15	G.COM	-0,87	0,44	G.YTH	-0,76	0,17
			G.MU	-0,45	0,22			
			MULT	0,70	0,109			

Tabella 4.8: ACM sulla matrice delle coproduzioni teatrali 2011-2012. Inerzia: 45,63% (non rival.). Organizzazione e Coproduzioni rilevanti per le prime tre dimensioni.

stati inseriti nell'analisi, rispettivamente, come variabili e punti supplementari. Analizzando i principali risultati dell'analisi riportati nella Tabella 4.8 e nelle mappe fattoriali (4.10 e 4.11) è possibile estrarre alcune informazioni. Prima di tutto, si riscontra una suddivisione in gruppi più complessa ma molto simile a quella riscontrata nell'analisi della matrice ridotta. Le prime tre dimensioni considerate racchiudono il 45,63% dell'inerzia (non rival.) complessiva. Ogni gruppo di organizzazioni ed eventi concorre alla caratterizzazione di una dimensioni o di una parte di essa facendo emergere distinti *pattern* di collaborazione. La prima dimensione è caratterizzata principalmente dalle compagnie Stabile Napoli e Campania Festival. Sulla seconda dimensione, si riscontra la contrapposizione tra Teatro Azione e Magazzini, da un lato, e Totò e Bracco, dall'altro che presentano modelli di collaborazione in opposizione. La terza dimensione è caratterizzata da Nuvole e ProgMuseo.

In particolare, dalle mappe fattoriali, grazie alla presenza degli attributi (delle compagnie e delle coproduzioni), si possono apprezzare le specificità dei modelli di connessione. Sulla mappa delle compagnie, rappresentate nello spazio del-

4.2 Analisi delle Corrispondenze Multiple e Blockmodeling Pre-specificato

le coproduzioni, (4.10) si può notare come quelle con *pattern* di collaborazione più comune, e in questo caso caratterizzato poche collaborazioni, si collochino nel centro della mappa, in prossimità dell'origine degli assi. Diversamente, le compagnie con *pattern* di collaborazioni inusuali, con la partecipazione a molte coproduzioni, si collochino agli estremi della mappa. Da quest'ultima emergono chiaramente almeno tre *pattern* distinti e in opposizione tra loro. Il primo è dato da Stabile Napoli e Campania Festival; essi hanno un importante ruolo economico (*producing house*, ossia teatro stabile riconosciuto a livello statale e finanziato) ed organizzativo (festival teatrali) nel sistema teatrale campano e sono coinvolte in coproduzioni caratterizzate dal genere classico, con script delle opere rivisti (ossia reinterpretati). Il secondo *pattern*, collocato nella parte alta del secondo asse, è dato da Teatro Azione e Magazzini (*building based companies*, ossia impresa di produzione teatrale) coinvolti in coproduzioni di genere classico, contemporaneo, storico biografico, caratterizzate da multidisciplinarietà, legati a spazi teatrali per le rappresentazioni di medie dimensioni. Il terzo *pattern*, collocato nella parte bassa del secondo asse, in diretta opposizione al precedente, è dato da Totò e Bracco (entrambe, *building based companies*), coinvolti in coproduzioni in prevalenza costituite da commedia e *musical* caratterizzate da script non rivisti (semplice riproduzione), legati a spazi teatrali di grandi dimensioni. A questi *pattern* se ne aggiunge un quarto, quello dato da Nuvole e ProgMuseo, (rispettivamente, *producing house* e istituzione educativo e culturale) nella parte bassa della quarta dimensione con coproduzioni caratterizzate dal genere del teatro per ragazzi e multidisciplinarietà, legati a spazi teatrali per le rappresentazioni di piccole dimensioni. Queste osservazioni trovano riscontro anche nella mappa fattoriale (4.11) che rappresenta le coproduzioni nello spazio delle compagnie. Con un primo gruppo di coproduzioni nel quadrante destro del primo asse fattoriale; due gruppi ben distinti e in opposizione sul secondo asse; un quarto gruppo sul quadrante inferiore del terzo asse della soluzione fattoriale.

Dalle osservazioni sin ora esposte si può comprendere in che modo tali informazioni possano essere utili in termini di pre-specificazione.

Prima di tutto, si possono trarre informazioni sul numero di cluster. Dalla struttura del network emergono almeno quattro *pattern* di relazioni distinti, quindi, in termini di Blockmodeling questo si traduce in un two-clustering composto da

4.2 Analisi delle Corrispondenze Multiple e Blockmodeling Pre-specificato

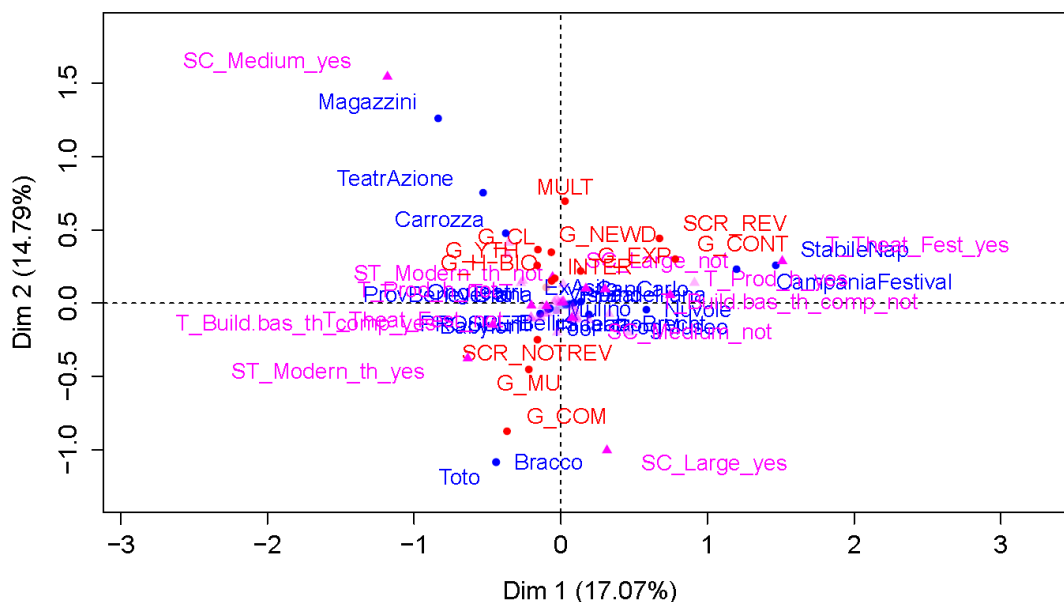


Figura 4.10: Mappa fattoriale delle coproduzioni teatrali. Compagnie nello spazio delle coproduzioni. Primo e secondo asse; Inerzia: 31,87% (non rival.)

almeno quattro cluster per le compagnie e quattro cluster per le coproduzioni, con l'aggiunta di un quinto cluster per le compagnie che hanno coproduzioni sparse. In secondo luogo, dato che ognuno dei quattro gruppi di compagnie si distingue in modo netto da ogni altro è possibile immaginare uno schema di vincoli alla struttura del clustering delle stesse in modo che in ognuno dei cluster siano presenti le compagnie individuate sul piano fattoriale. Ad esempio si può imporre che Stabile Napoli e Campania Festival appartengano allo stesso cluster, lo stesso per gli altri gruppi.

In terzo luogo è possibile immaginare i tipi di blocchi ideali più adatti a descrivere i modelli di connessioni individuati. Il *pattern* di collaborazione di Stabile Napoli e Campania Festival può essere descritto da un blocco ideale completo, i punti che definiscono i profili di partecipazione dei due attori occupano una collocazione nello spazio molto simile, sono distribuzionalmente equivalenti, suggerendo una quasi completa condivisione delle coproduzioni in cui sono coinvolte. Lo stesso vale per Totò e Bracco, i punti che indicano le due compagnie sono indistinguibili

4.2 Analisi delle Corrispondenze Multiple e Blockmodeling Pre-specificato

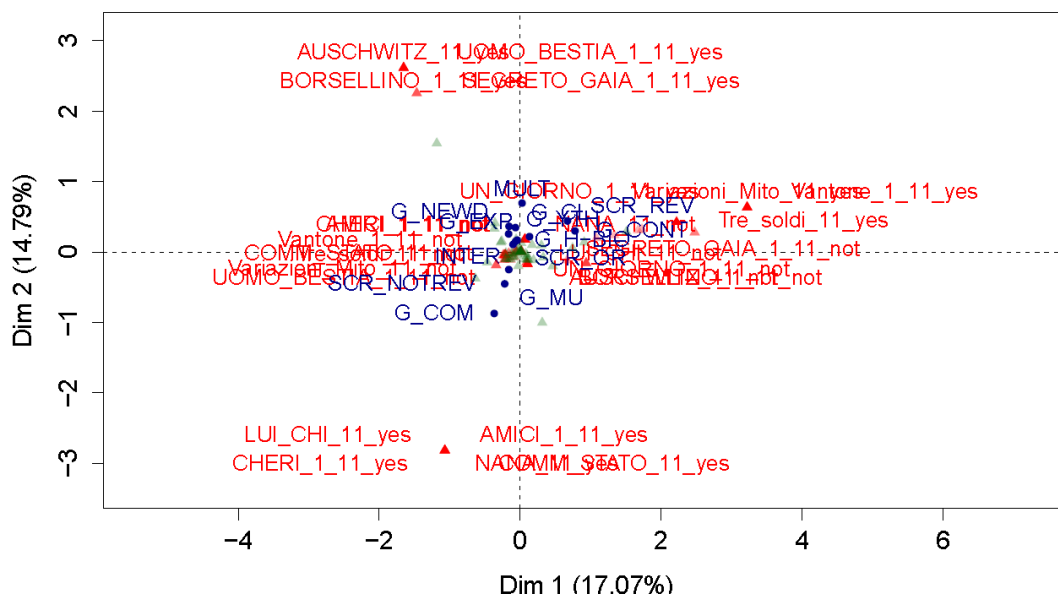


Figura 4.11: Mappa fattoriale delle coproduzioni teatrali. Coproduzioni nello spazio delle compagnie. Primo e secondo asse; Inerzia: 31,87% (non rival.)

indicando una perfetta sovrapposizione dei profili di partecipazione. Il *pattern* di collaborazione di Nuvoles e ProgMuseo è configurabile come completo o denso, i profili si collaborano sono molto vicini indicando la condivisione di buona parte delle coproduzioni. Anche il modello di connessioni di Teatro Azione e Magazzini suggerisce un blocco denso, con la condivisione delle maggior parte delle coproduzioni. Per il modello di connessioni delle compagnie con coproduzioni sparse non sembra possibile individuare un blocco ideale in anticipo come pure non è detto che costituisca un unico cluster. In ultimo è possibile definire un modello di struttura del blockmodel. Considerando che ad ogni gruppo di compagnie nello spazio delle coproduzioni, sembra corrispondere un gruppo di coproduzioni nello spazio delle compagnie, è possibile definire una struttura segmentata, con i blocchi indicati collocati lungo la diagonale della struttura del modello.

Le indicazioni così definite possono essere utilizzate per definire un modello di struttura di clustering e/o di blockmodel basato su informazioni relazionali e descrittive empiriche. Tale modello di pre-specificazione può essere "testato" attraverso un'analisi di Blockmodeling. I vantaggi di una strategia simili riguardano

4.2 Analisi delle Corrispondenze Multiple e Blockmodeling Pre-specificato

sia la riduzione del numero di tentativi necessari all'individuazione del numero di cluster e del blockmodel ottimale, sia la possibilità di avere una base di conoscenze sulla quale fondare la scelta tra partizioni egualmente ben modellate, con valori delle funzione criterio minimi. Le informazioni sono state ottenute attraverso la ricostruzione dei modelli relazionali e del contesto in cui le relazioni sono calate e dal quale possono essere influenzate. In questo modo, partizioni analoghe, rispetto ad un certo criterio di equivalenza, possono essere confrontate non solo in base a considerazioni di carattere teorico ma anche empirico. Inoltre, è necessario sottolineare come una strategia di analisi come quella proposta in questo lavoro implica un continuo scambio tra una fase esplorativa e una confermativa privilegiando un approccio di "scoperta e costruzione continua della conoscenza" simile a quello adottato nell'ambito del *Knowledge Discovery Process* (2, 137). In questo modo, la produzione della conoscenza segue un processo circolare guidato dai dati la cui esplorazione consente l'individuazione di *pattern* di relazioni rilevanti. A partire da tali informazioni è possibile formulare delle ipotesi sulla struttura di dati, le quali, a loro volta, possono essere sottoposte a verifica attraverso un approccio confermativo. Il risultato di tale processo fornisce ulteriori intuizioni sul fenomeno in esame, utili a formulare nuove ipotesi e dare origine ad un nuovo ciclo. Nel presente lavoro, proponiamo di adottare un approccio basato su una interazione continua tra l'aspetto induttivo e quello deduttivo, che parte dall'esplorazione dei *pattern* di relazioni attraverso l'AF che consente di definire un'ipotesi sul modello di struttura esistente nei dati da testare, e quindi verificare, attraverso l'analisi di Blockmodeling. Allo stesso tempo, i risultati del Blockmodeling consentono di riformulare e ri-specificare il modello di struttura più adeguato ai modelli di connessioni e, a loro volta, i risultati che ne derivano possono essere ulteriormente introdotti in una cornice di analisi multidimensionale. In questo modo essi vengono analizzati e rappresentati in relazione ad altre informazioni che ne consentono una migliore interpretazione.

4.3 Analisi delle Corrispondenze Multiple e Blockmodeling: combinazione

L'Analisi delle Corrispondenze Multiple offre la possibilità di integrare al suo interno, come variabili e punti supplementari, gli stessi risultati del Blockmodeling. Di conseguenza, il blockmodel può essere messo in relazione alla struttura dei legami sottostanti, e alle altre informazioni descrittive e strutturali, sia dal punto di vista analitico sia visuale. In effetti, uno dei vantaggi più interessanti è dato dal fatto che le tecniche di Analisi Fattoriale consentono la proiezione delle posizioni, degli attori e degli eventi, in uno spazio metrico che rende possibile una rappresentazione pienamente significativa, in cui lo spazio tra le posizioni può essere interpretato in termini di distanze. Il tema della rappresentazione dei dati relazionali in uno spazio metrico è stato sempre di grande interesse nella disciplina della SNA, a partire dagli sviluppi delle tecniche di Scaling Multidimensionale e dall'applicazione dell'Analisi in Componenti Principali e di varie versioni dell'AC e dell'Analisi Fattoriale in generale (105). Inoltre, spesso è stato sottolineato che il modo in cui i dati vengono rappresentati può influenzare la percezione delle caratteristiche strutturali del network (112, 148). Si riscontra che le visualizzazioni basate sulle distanze euclidee sono quelle ad essere meno fuorvianti. In effetti, le rappresentazioni ottenute attraverso i grafi, non rispettando le regole della geometria euclidea, non consentono di valutare lo spazio tra i nodi come distanza. In questo contesto, quindi, si propone di incorporare i risultati del blockmodel in uno spazio metrico costruito sulla base degli stessi dati relazionali che consenta di valutare la collocazione e le distanze delle posizioni nello spazio relazionale. L'opportunità di inserire le posizioni in un contesto informativo più ampio ne consente la caratterizzazione ed una interpretazione più profonda. Pertanto, la combinazione di tecniche di Analisi Fattoriale e di Blockmodeling offre la possibilità di valorizzare i singoli contributi informativi e migliorare l'interpretazione globale dei risultati con riflessi positivi sulla capacità di comprensione del fenomeno sociale in esame.

4.3.1 Visualizzazione e caratterizzazione delle posizioni

Nel corso di quest'ultima parte viene illustrato uno dei risultati relativi all'analisi della collaborazione nelle coproduzioni teatrali in quattro stagioni teatrali (che costituisce un *time-varying two-mode network*), realizzata attraverso l'Analisi Fattoriale Multipla (AFM). I dettagli del metodo e dell'analisi verranno forniti nel prossimo capitolo specificatamente dedicato ad esso. In quest'ultima parte viene anticipata la rappresentazione grafica delle compagnie, nello spazio degli eventi (4.12). Nella mappa fattoriale sono stati proiettati anche gli attributi descrittivi e le posizioni degli eventi (come punti supplementari). L'introduzione delle posizioni in uno spazio metrico consente di esplorare l'associazione tra i modelli relazionali e gli attributi che li caratterizzano e, nel caso dell'AFM anche la loro evoluzione temporale. La mappa fattoriale costruita a partire dalle prime due dimensioni relative all'analisi globale riproduce il 27,41% (non rival.) dell'inerzia complessiva. Su di essa è possibile analizzare cumulativamente i quattro istanti temporali, cercando la struttura relazionale complessiva racchiusa nella matrice di affiliazione globale e valutando la collocazione delle posizioni degli eventi/attori in relazione agli attori/eventi a cui sono associati e rispetto alle loro caratteristiche descrittive; evidenziare le differenze tra gli istanti temporali in termini di struttura globale, e analizzare le variazioni degli attori/eventi e della rispettive posizioni di eventi/attori a cui sono associati (nel tempo). Questi ultimi due punti permettono di studiare la variazione longitudinale e il comportamento nel tempo dei *pattern* relazionali degli attori/eventi e delle posizioni a cui sono associati. Le visualizzazioni grafiche così ottenute, consentono di collocare le posizioni in uno spazio metrico e sostantivo che ne facilita e migliora la caratterizzazione e l'interpretazione.

4.4 Analisi delle Corrispondenze Multiple e Blockmodeling: vantaggi

In conclusione, è possibile evidenziare alcuni punti della nostra proposta metodologica che sembrano particolarmente significativi.

4.4 Analisi delle Corrispondenze Multiple e Blockmodeling: vantaggi

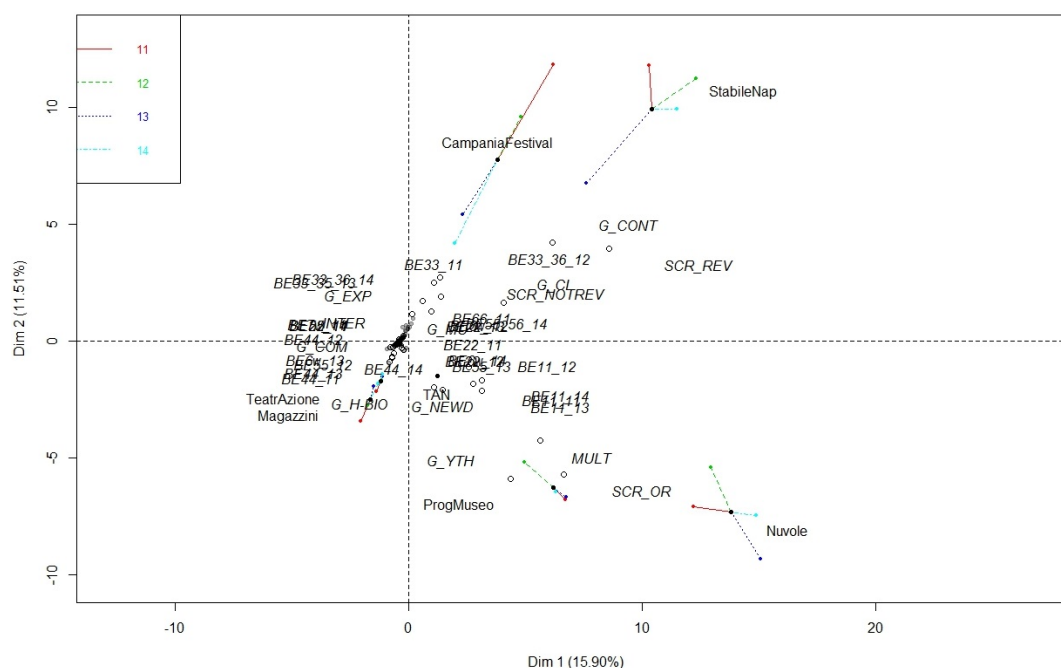


Figura 4.12: Mappa fattoriale delle coproduzioni teatrali. Compagnie e posizioni delle coproduzioni nello spazio delle coproduzioni. Primo e secondo asse; Inerzia: 27,41% (non rival.)

Nello specifico, si rileva come la procedura di ricostruzione della matrice originaria operata attraverso il metodo di Eckart-Young consenta di ripulire la struttura del network, conservando i *pattern* più rilevanti. Infatti, l'operazione di "filtraggio", operata dalla ricostruzione attraverso i primi valori singolari, riduce quelle che nel Blockmodeling potrebbero essere di inconsistenze poiché ha l'effetto di "regolarizzare i *pattern* relazionali". Tale operazione, considerata la struttura generale, tende ad eliminare i legami dove non dovrebbero essere presenti (che, nel Blockmodeling, corrispondono alla presenza di legami nei blocchi che dovrebbero essere nulli), e tende ad aggiungerli dove dovrebbero esserci (che, nel Blockmodeling, corrispondono all'assenza di legami nei blocchi che dovrebbero essere completi). In termini di Blockmodeling, questo significa proprio eliminare legami presenti nei blocchi che dovrebbero essere nulli ed aggiungere legami nei blocchi che dovrebbero essere completi (sempre, in base alla struttura generale

del network). L'operazione di ricostruzione basata sul metodo di Eckart-Young tende a creare dei blocchi omogenei, nulli o completi, seguendo la stessa logica alla base del Blockmodeling eseguito in base al criterio dell'equivalenza strutturale. Ad esempio, se nella matrice originaria due attori sono legati insieme ad un certo numero di eventi di un sottoinsieme, ma non precisamente a tutti, ad esempio 4 su 5, la procedura di ricostruzione tenderà ad assegnare un legame all'attore anche per quell'evento non condiviso, restituendo un *pattern* di connessioni omogeneo. Riteniamo che questa caratteristica della procedura di ricostruzione la renda particolarmente utile nell'operazione di ripulitura e di sintesi della struttura "forte" come fase di pre-trattamento preliminare nell'analisi di Blockmodeling Esplorativo. In quest'ottica, tale strategia è in linea con gli approcci di analisi statistica multidimensionale che mostrano come la capacità di recupero delle informazioni delle procedure di clustering possa essere migliorata, e resa meno sensibile alla presenza di rumore nei dati, utilizzando i metodi fattoriali prima della *Cluster Analysis* (19, 138). Tale approccio non è esente da critiche (172, 188), tuttavia, riteniamo che costituisca un primo passo verso la definizione di una strategia che preveda una vera e propria integrazione dei due metodi.

In aggiunta, è necessario evidenziare che la rappresentazione grafica ottenuta attraverso i primi assi fattoriali, tramite l'ACM, risulta coerente con la logica dell'equivalenza strutturale e ciò fornisce un'ulteriore legittimazione per l'utilizzo di tale metodo sia per la ricostruzione di informazioni utili nel Blockmodeling Pre-specificato sia per la rappresentazione dei risultati del Blockmodeling attraverso tale metodo.

In sintesi, si può concludere che l'ausilio delle tecniche Fattoriali nell'analisi del *two-mode* Blockmodeling consente: la sintesi della struttura del network, ripulendola da una parte di elementi non caratterizzanti; l'individuazione di un modello di struttura per la pre-specificazione; fornisce un sottospazio di riferimento metrico in cui localizzare le posizioni in base ai *pattern* di similarità nei profili di partecipazione degli attori ed eventi che ne fanno parte. In questo contesto, l'applicazione delle tecniche fattoriali al trattamento dei network *two-mode* si pone come strumento analitico e visuale in grado di sintetizzare la struttura sottostante e non solo come metodo di scaling e di visualizzazione.

Capitolo 5

Analisi Fattoriale Multipla e Blockmodeling: un approccio combinato per l'analisi del campo

5.1 Analisi Fattoriale Multipla per l'analisi di Network *Two-Mode* eterogenei

Una delle problematiche che sorgono dal dibattito scientifico nell'ambito della SNA è legato all'analisi dei network eterogenei (149, 185). La maggior parte dei metodi esistenti sono stati sviluppati per i network *one-mode* (solo in parte per i network *two-mode*), ad una sola relazione, prevalentemente sulla base delle informazioni relative ai legami. Tuttavia, un network sociale può assumere una struttura molto complessa in cui, oltre alla dimensione, possono esservi più di una classe di entità (*multi-mode*), più di una relazione (*multi-relational*). Inoltre, possono essere disponibili molteplici tipi di informazioni, associate sia alle entità sia alle relazioni, come attributi descrittivi e proprietà strutturali, le quali, a loro volta, possono riguardare diversi momenti temporali (149, 185). Per l'analisi dei network eterogenei è necessario dotarsi di strumenti flessibili che consentano di combinare i vari tipi di informazioni contenuti in essi e di estrarre la struttura relazionale sottostante. In questo lavoro proponiamo un approccio per analizzare ed esplorare visivamente network eterogenei. Nello specifico ci si riferisce a *two-mode* network, ad una sola relazione rilevata in momenti temporali diversi (occasioni), con attributi descrittivi e proprietà strutturali come informazioni aggiuntive. Le

5.1 Analisi Fattoriale Multipla per l'analisi di Network *Two-Mode* eterogenei

opportunità derivanti dalla combinazione di tecniche di Analisi Fattoriale e di Blockmodeling sono visibili anche in una prospettiva più ampia legata alla gestione di network eterogenei. I fenomeni sociali sono complessi, multidimensionali e pertanto la loro comprensione richiede la capacità di mettere in relazione aspetti diversi a differenti livelli di analisi. Ne consegue la necessità di trattare informazioni molteplici e di natura diversa, spesso, in un'ottica diacronica. Nel tentativo di compiere un passo in questa direzione proponiamo una strategia di analisi che combina l'Analisi Fattoriale Multipla (AFM)¹ (88, 164, 169) con il Blockmodeling per lo studio dei network *two-mode* eterogenei, in cui sono presenti dati relativi ai legami tra due classi di entità differenti, in momenti temporali diversi, con attributi descritti e proprietà e i modelli di relazioni che intercorrono tra loro. Questa strategia, pur non essendo in grado di cogliere pienamente la dinamica processuale alla base dell'evoluzione della struttura, consente di esplorare, descrivere e rappresentarne in modo accurato gli esiti mettendoli in relazione reciproca. Tale approccio analitico è stato adottato per lo studio del campo della produzione teatrale nell'ambito del sistema teatrale regionale campano (182). Nello specifico, lo studio è volto a cogliere la struttura del campo delle co-produzioni teatrali attraverso una integrazione della Teoria dei Campi di Pierre Bourdieu (35, 36) e dell'Approccio Posizionale (54, 56, 83) della SNA. Tale scelta nel tentativo di evidenziare come l'approccio multidimensionale, privilegiato da Bourdieu per il suo carattere "relazionale" capace di portare alla luce le "relazioni oggettive", e quello della SNA, "scartato" dallo stesso per esser eccessivamente ingessato nella descrizione di particolari legami piuttosto che alla struttura sottostante, possano essere non solo compatibili ma anche fruttuosamente integrabili. Il punto di partenza della connessione tra "struttura relazionale del network" e "struttura oggettiva" è legato al fatto che entrambe sono considerate come parte di un campo, in generale, e, quindi, anche del campo della produzione culturale. Riteniamo che l'approccio posizionale della SNA sia in grado di rivelare la struttura di relazioni che sostiene quelle oggettive e che, contemporaneamente, costituisce il *frame* in cui essa prende forma (183). In letteratura non sono disponibili molti strumenti

¹Per un breve dettaglio tecnico si rinvia all'appendice 5.5, per una trattazione dettagliata sul metodo dell'AFM per network *two-mode time-varying* si rinvia a Ragozini et al. (169).

analitici per il trattamento di network *two-mode* che variano nel tempo (128), tuttavia, diversi sono quelli disponibili per la visualizzazione dei dati relazionali che consentono rappresentazioni sia statiche sia dinamiche (11, 42, 104). I network *two-mode time-varying*, possono essere organizzati in matrice multidimensionali, in questo modo la struttura sottostante i dati può essere analizzata attraverso i metodi statistici per l'analisi dei dati di tabelle a più vie (*multiway data analysis*) (134). Tra i metodi esistenti, in questo lavoro, viene adottato l'Analisi Fattoriale Multipla (AFM)(87, 164), il cui utilizzo per il trattamento dei network *two-mode time-varying* sembra particolarmente adeguato(169). L'AFM è un'estensione delle tecniche fattoriali (88, 164) strutturata per la gestione di tabelle di dati multiple che possono includere sia le stesse osservazioni su diversi insiemi di variabili sia le stesse variabili misurate su diversi set di osservazioni. Questo rende possibile l'analisi congiunta di variabili sia quantitative sia qualitative, fornendo visualizzazioni in cui le rappresentazioni del set di individui associati a ciascun gruppo di variabili vengono sovrapposti. In questo modo, l'AFM si pone come uno degli strumenti più adeguati a gestire l'eterogeneità dei fenomeni complessi. L'AFM, nel presente lavoro, viene applicata seguendo un approccio diretto alla matrice di affiliazione (30, 71). Questo approccio consente di ottenere delle rappresentazioni grafiche che racchiudono sia la dimensione temporale sia la descrizione analitica del network che permettono delle visualizzazioni grafiche, che pur essendo statiche, danno la possibilità di descrivere l'evoluzione del network nel corso del tempo, preservando le diverse caratteristiche dei due modi (169).

5.2 Posizioni nella *Social Network Analysis* e nella Teoria del Campo

Un approccio molto interessante nell'analisi dei fenomeni sociali che presenta affinità concettuali con l'analisi posizionale nella SNA, che può essere applicato anche allo studio della produzione culturale, è dato dalla teoria del campo di Pierre Bourdieu (35). Quest'ultimo definisce il campo come un network di relazioni, o meglio, come un network, o una configurazione, di relazioni oggettive tra posizioni. La struttura del campo, ossia lo spazio delle posizioni, è generato dalla struttura della distribuzione di specifiche risorse, o forme di capitale, che

regolano e determinano il successo nel campo. Inoltre, ogni posizione è interrelata alle altre posizioni presenti nel campo, sia per la sua esistenza sia per gli effetti che genera per chi le occupa (35). Nella teoria del campo, quindi, il concetto di posizione è legato alla collocazione dei soggetti in uno spazio sociale in virtù della loro relazione rispetto alle risorse rilevanti presenti nel campo. I soggetti che occupano la stessa posizione sono in una relazione di equivalenza poichè equivalente è il loro rapporto con le risorse e, quindi, con le altre posizioni.

Come si è visto nel capitolo secondo, nell'approccio posizionale il concetto di posizione si riferisce ad un insieme di individui che sono incorporati allo stesso modo nel network di relazioni, mentre il ruolo si riferisce ai modelli di relazioni che intercorrono tra gli attori o tra le posizioni. Questa definizione richiama il concetto fondante di equivalenza strutturale, per cui gli attori sono strutturalmente equivalenti se hanno lo stesso modello di connessioni con gli stessi attori. Di conseguenza, l'equivalenza posizionale, in termini di network, implica che gli attori siano stati assegnati ad un dato insieme di classi di equivalenza sulla base dei loro modelli di connessioni. Come si è ampiamente discusso 2.2, una degli approcci più adatti ad analizzare questo tipo di proprietà è quello del Block-modeling (77, 83, 98), che, pertanto, è stato utilizzato anche dagli studiosi che hanno applicato la teoria dei campo nell'ambito della SNA (34, 110). Gli studiosi che adottano questa scelta ritengono che l'analisi posizionale sia adatta ad operativizzare parte della concezione della teoria dei campi di Bourdieu. L'idea che attori occupanti la stessa posizione, quindi strutturalmente equivalenti, siano simili nel loro modello di connessioni e non necessariamente sono in contatto tra loro, fa sì che occupino una posizione in uno spazio sociale astratto (183). Pertanto, nell'approccio posizionale, le relazioni si pongono a un livello di astrazione che produce una visione oggettiva, che permette al ricercatore di comprendere il significato delle posizioni in una struttura sociale indipendentemente dai loro specifici occupanti (183). La struttura di un blockmodel è sia una rappresentazione del network osservato sia un'astrazione strutturale di un network di dati (22, 198). Inoltre, come ulteriore punto di connessione con l'approccio di Bourdieu, l'equivalenza posizionale ha delle implicazioni anche rispetto alla struttura del network, alla disponibilità delle risorse e anche alla cultura, in quanto alle posizioni sono legate anche i vincoli e le opportunità. Queste ultime, sono ciò

che differenzia le posizioni nella struttura del campo in termini di potere, prestigio, ecc. In aggiunta, i vantaggi e gli svantaggi sono associati alla distribuzione delle varie forme di capitale che possono transitare attraverso i legami sociali. Pertanto, nonostante la "struttura oggettiva" riguardi esclusivamente la distribuzione delle varie forme di capitale nello spazio sociale, questo è inevitabilmente associato ad un certo insieme di *pattern* relazionali che possono essere delineati attraverso l'analisi posizionale (183).

5.3 Network e campo

Bourdieu definisce i campi come configurazioni di relazioni oggettive tra posizioni ma anche come luoghi di competizioni strategiche, materiali e simboliche, tra attori che dispongono di vari gradi e forme di capitale (culturale, economico, sociale, simbolico) (35). Nella sua visione, i contenuti e la composizione delle risorse di cui si dispone influenzano anche le relazioni sociali che esso può intrattenere.

L'industria del teatro può essere concepita come un sistema organizzato in modo relazionale in cui i "prodotti culturali sono generati in e da network di collaborazione tra imprese e tra persone" (74, p. 287). Tuttavia, nonostante la presenza notevole di studi che adottano una prospettiva relazionale attraverso la SNA in vari ambiti della produzione artistica (4, 65, 67, 114, 147), sono relativamente pochi gli studi sui network culturali che affrontano l'analisi della produzione teatrale nella sua dinamica e nella sua struttura relazionale (73).

La produzione teatrale come sistema relazionale può essere pensato come un *campo* della produzione culturale (183) in modo simile a come Bourdieu (35, 40) definisce il campo della letteratura e, più in generale, il campo artistico: una network o uno "spazio delle relazioni oggettive tra le posizioni" (40, p. 181). In questo modo, è possibile cogliere le competizioni sottostanti tra produttori culturali per occupare e mantenere posizioni importanti in uno spazio sociale definito dalla distribuzione delle varie forme di risorse (183). L'approccio di Bourdieu, tuttavia, è orientato alla definizione delle relazioni strutturali in uno spazio sociale costruito analiticamente a partire dagli attributi categoriali (183), e mette in discussione la SNA come prospettiva in grado di cogliere tali relazioni (38, 41). Malgrado ciò, l'approccio di Bourdieu non è necessariamente inconciliabile con

l'analisi delle relazioni concrete tra gli attori che occupano posizioni all'interno dei campi, realizzabile attraverso la SNA (34, 68, 86).

L'obiettivo del presente studio di caso è dato dal tentativo di esplorare e analizzare un campo concepito come uno spazio sociale adottando un approccio concettuale e metodologico che consenta di tenere insieme sia la struttura relazionale, legata alle relazioni sociali, sia quella oggettiva, legata alla distribuzione delle varie forme di risorse che lo caratterizzano. Nell'approccio da noi proposto, si sceglie di analizzare il campo della produzione teatrale, tuttavia, tale approccio può essere applicato anche allo studio di altri fenomeni sociali. In questo studio di caso il campo della produzione teatrale viene concepito come uno spazio di relazioni inter-organizzative. Nel nostro approccio lo studio della struttura del campo parte dall'analisi di un network *two-mode time varying* costituita da diverse co-produzioni teatrali che coinvolgono i teatri e le compagnie teatrali situate nella regione Campania, Italia, per quattro stagioni teatrali (181, 183). L'analisi di tale campo organizzativo e culturale si concentra sui (i) modelli relazionali e le caratteristiche sia delle compagnie che delle co-produzioni; (ii) le gerarchie e la segmentazione della struttura relazionale; (iii) la dinamica dei modelli relazionali nel tempo. Tale analisi si basa sull'adozione della prospettiva della *teoria dei campi* di Bourdieu (35, 40), che offre strumenti concettuali applicabili anche per lo studio del contesto organizzativo nel settore artistico, consentendo l'integrazione sia della struttura relazionale sia della struttura oggettiva in un unico *framework* concettuale e approccio metodologico (183). L'approccio metodologico e statistico adottato, in parte, si iscrive nell'ambito dell'analisi posizionale della SNA (54, 56) e si basa sulla combinazione delle tecniche di Blockmodeling e di Analisi Fattoriale. Nello specifico, l'analisi dei network *two-mode* delle co-produzioni teatrali è condotta attraverso il *two-mode* Blockmodeling (81, 83) e l'AFM ((88, 163, 169). La prima tecnica consente di estrarre le posizioni strutturali, in termini di relazioni di network, la seconda consente di inserire queste ultime in uno spazio sociale definito dalle relazioni sociali e dagli attributi categoriali sottostanti (ossia, dalla distribuzione delle varie forme di risorse). L'AFM prevede l'applicazione di una tecnica statistica molto cara a Bourdieu nell'analisi delle "relazioni oggettive", ossia l'ACM, e consente un'analisi multidimensionale

in grado di tenere conto anche della dimensione diacronica e, quindi, della variazione della struttura del network nel tempo. Tale approccio, nel suo insieme, permette lo studio della struttura del campo della produzione teatrale, racchiusa nel network *two-mode*, definendo, da un lato, le posizioni occupate dalle compagnie nello spazio sociale e, dall'altro, le loro "prese di posizioni", ossia le scelte manifeste esercitate dalle compagnie in virtù dell'appartenenza a determinate posizioni, e che si esprimono attraverso specifiche scelte di co-produzione.

5.4 Analisi del campo attraverso l'Analisi Fattoriale Multipla e il Blockmodeling: il caso della produzione teatrale

Considerato quanto emerso sino ad ora, la struttura oggettiva e la struttura relazionale possono essere viste come strettamente interrelate. In quest'ottica, nel presente lavoro, proponiamo un approccio orientato all'inclusione delle relazioni di network nell'analisi del campo in modo da descrivere e analizzare la struttura oggettiva in modo più profondo. Pertanto, tenere conto delle relazioni intersoggettive - che, in questo caso, sono i rapporti inter-organizzativi - può essere utile per far emergere differenze o somiglianze tra posizioni e prese di posizione nello spazio sociale. Questo approccio è orientato a far emergere le relazioni oggettive dai network di relazioni osservate (34) e fornisce un modo per ricostruire la struttura del campo (68). In questo modo, i ruoli (come dominanti *versus* dominati) possono essere derivate dal tipo specifico di posizioni che emergono dall'analisi (183). Nel tentativo di seguire questa strategia si adotta l'analisi posizionale attraverso il Blockmodeling, insieme all'AFM, che, di base include l'utilizzo dell'MCA come metodo per esplorare le relazioni esistenti nei dati. L'AC e l'ACM (20, 119) sono state le tecniche privilegiate da Bourdieu nell'analisi delle "topologia sociale" (36, 37). Nella logica dell'AC, lo spazio sociale, e nel nostro caso dei produttori culturali, può essere analizzato come uno spazio geografico diviso in regioni (38), sulla base alle loro caratteristiche legate agli indicatori delle forme di capitale, mentre il Blockmodeling fornisce un adeguato partizionamento della struttura del network. Tale approccio, che utilizza congiuntamente le due tecniche, è stato adottato in letteratura per formalizzare e testare il concetto di campo della produzione culturale nel mondo letterario (4, 110). Le ragioni alla base della

5.4 Analisi del campo attraverso l'Analisi Fattoriale Multipla e il Blockmodeling: il caso della produzione teatrale

scelta di questo approccio sono legate alla convinzione che la struttura di network e la struttura oggettiva possano essere uniti in un unico *framework* concettuale in cui il duplice uso della nozione di posizione (nella SNA e nella teoria dei campi) viene perseguito attraverso l'analisi multidimensionale che si presenta come l'approccio più appropriato dati i presupposti concettuali della teoria dei campi e dell'approccio della SNA ai campi della produzione culturale. In questo modo, la SNA e l'AC grazie alle loro specificità tecniche consentono di mappare la struttura del campo. In tal senso, l'obiettivo del presente lavoro si traduce nell'utilizzo dell'analisi posizionale attraverso l'ACM (inclusa nell'AFM) per mappare al suo interno le posizioni equivalenti insieme agli attributi categoriali. A tal fine, proponiamo un approccio combinato basato sull'MCA e Blockmodeling allo scopo di analizzare il campo sia in termini di *pattern* di relazioni sia di distribuzione delle diverse forme di capitale. La novità della nostra proposta, rispetto alle altre applicazioni del concetto di campo nella SNA, risiede nel modo in cui questi metodi vengono usati.

Il punto di partenza fondamentale è la connessione tra la struttura di network e la struttura oggettiva, entrambe sono considerate come parti costitutive di un campo, che nel presente studio di caso è dato da quello della di produzione culturale. Si sostiene che l'approccio posizionale della SNA possa rivelare la struttura relazionale che sostiene quella oggettiva. In questo senso la struttura relazionale giace al di sotto di quella oggettiva e costituisce una sorta di *frame* in cui essa prende forma. Pertanto, la struttura di network e la struttura oggettiva sono parallele tra loro. La distribuzione complessiva del capitale e quella delle relazioni di network operano simultaneamente e implicano l'esistenza l'una dell'altra (183). Infatti, le relazioni inter-soggettive possono mediare e influire su quelle oggettive (68).

Di conseguenza, l'obiettivo del presente studio si traduce nell'analisi della struttura del campo, definito dalla sovrapposizione di due strutture di natura diversa. Da un lato, la struttura oggettiva data dalla distribuzione delle risorse (misurate attraverso specifiche variabili categoriali, ossia gli attributi) che determina lo spazio delle posizioni oggettive. Dall'altro la struttura relazionale data dai *pattern* di connessioni che determina le posizioni strutturali. Questa operazione, nel presente caso, diviene più complessa dal considerare un network *two-mode*

che implica il concetto di dualità (43) rintracciabile anche nella teoria dei campi. In questo caso, infatti, gli attributi e le relazioni sono considerati rispetto sia ai produttori sia ai prodotti culturali. Questa strategia di ricerca è mossa dal tentativo di tenere insieme la dimensione teorica e quella empirica, rilevante sia nell'ambito della ricerca sui campi sia nella SNA (86, 183).

5.4.1 Sistema teatrale campano: network delle co-produzioni come network *two-mode*

Quando formalizzata in termini di network, la maggior parte dei campi della produzione culturale possono essere concepiti come sistemi strutturati in modo duale costituiti dai creatori e dai loro prodotti creativi, nel presente caso attori e produzioni teatrali. L'oggetto del presente lavoro è il legame esistente tra le organizzazioni creativi e dei loro prodotti, ossia, compagnie teatrali e produzioni teatrali. Ognuna delle due parti della connessione può essere rappresentata come un insieme di unità sociali o culturali in un network *two-mode* e, nello specifico, in network di affiliazione (30, 192) in cui le organizzazioni e i prodotti sono legati tramite la partecipazione del primo nella creazione del secondo (183). Questa struttura di dati consente di identificare la dimensione relazionale della struttura del campo, in quanto, si tratta di una struttura non diadica in cui le organizzazioni e prodotti creativi sono legati nella misura in cui sottogruppi di imprese sono legate alle opere che realizzano congiuntamente (93). In questo tipo di struttura i due insieme di entità sono legate reciprocamente racchiudendo la dimensione della dualità (43). Il termine "co-produzione" indica sia l'esito sia il processo stesso di un'attività basata su un progetto in cui due o più imprese cooperano, per un periodo di tempo limitato, in modo da produrre ed eseguire un'opera teatrale definendo una sorta di "sistema temporaneo" (115, 182, 183). Alla base della scelta di collaborazione possono essere rintracciate molteplici motivazioni, che vanno dalla condivisione di costi e risorse (finanziarie, ma anche umane), alla condivisione di un progetto artistico comune (comune visione estetica, rinnovare collaborazioni passate, volontà di lavorare con un certo partner, individuale o collettivo, per il suo prestigio o valore professionale). Le collaborazioni sono spesso associati con la complementarità di risorse e di interdipendenza. Pertanto, le varie forme di

5.4 Analisi del campo attraverso l'Analisi Fattoriale Multipla e il Blockmodeling: il caso della produzione teatrale

capitale, e la loro distribuzione disuguale all'interno di un campo, sono elementi importanti per la struttura del network (ad esempio, le disuguaglianze nei *pattern* relazionali o di capitale possono portare la struttura del campo, analizzata in termini di relazioni e attributi, ad assumere una certa forma coerente con un dato modello di connessioni, come il *core-periphery*, dovuta ad una concentrazione di potere o di opportunità in un certo punto della struttura)(183). Inoltre, in questo lavoro il termine 'co-produzione' è inteso in un senso molto ampio, che va al di là dell'accordo formale tra le organizzazioni. Infatti, si considerano diversi tipi di *partnership* formali e informali che, indipendentemente dal contratto, offrono la possibilità di stabilire o rafforzare legami tra le organizzazioni coinvolte nelle collaborazioni. Di conseguenza, le relazioni analizzate non riguardano esclusivamente le "interazioni" dirette tra i partecipanti in una serie di co-produzioni, ma anche le connessioni indirette che "connettono" i produttori culturali in virtù del loro legame con una serie di terze parti comuni che sono membri di alcune produzioni congiunte (183). Questo punto rappresenta un tratto essenziale della struttura di campo che verrà indagato. La condivisione, da parte di un'organizzazione, di uno o più co-produzioni con altre compagnie, può costituire una potenziale opportunità di contatto con altre organizzazioni legate alle compagnie partner, diversamente ne rimane esclusa. Questa osservazione porta a valutare nell'analisi il grado di inclusione/esclusione che caratterizza le organizzazioni nel network. Di un certo interesse è capire se, e in che misura, la struttura del network, rispetto alle organizzazioni teatrali o alle co-produzioni, tende ad assumere una forma segmentata in base alle loro caratteristiche socio-culturali. Un altro punto di potenziale interesse è legato al fatto che vengono considerate relazioni inter-organizzative a volte invisibili agli attori coinvolti in essi, al di là di un collegamento diretto tra i partecipanti nelle co-produzioni. Anche in questo caso, queste connessioni si discostano dalle interazioni "pure" non legate alla struttura oggettiva del campo (183). Nella misura in cui un certo numero di attori rilevanti possano esercitare un certo potere in termini di popolarità e di partecipazione alle co-produzioni, si pongono ulteriori interrogativi da tenere presenti nell'analisi, legati alle potenziali strutture di potere e alle gerarchie di classi di organizzazioni dominanti/dominati che possono venire a configurarsi (183). La natura duale della matrice di affiliazione consente di riprodurre la dimensione di un campo e

5.4 Analisi del campo attraverso l'Analisi Fattoriale Multipla e il Blockmodeling: il caso della produzione teatrale

le opposizioni tra specifiche categorie di produttori culturali così come tra i loro prodotti. Lo spazio sociale dei produttori di teatro è definito da una struttura duale in cui le distinzioni tra le produzioni caratterizzata da alcuni generi teatrali dovrebbe riflettere una distinzioni tra i produttori. Questa omologia viene riflessa nel network di affiliazione nella misura in cui i produttori sono affiliati gli uni agli altri attraverso i prodotti culturali(183). Ciò consente di concepire il network *two-mode* derivato dalle co-produzioni teatrali come un campo in cui le posizioni sono occupate dalle istituzioni coinvolte nelle co-produzioni. Pertanto, unendo la logica della dualità del network di affiliazione con quella della teoria dei campi, si può definire in primo set di nodi del network *two-mode*, ossia quello delle organizzazioni teatrali, come il set delle "posizioni" nella struttura di campo, mentre, il secondo set, dato dall'insieme delle eventi, o co-produzioni teatrali, come il set delle "prese di posizione" (183).

Data la dinamicità del campo, viene presa in considerazione anche la sua dimensione temporale, considerando le coproduzioni realizzate in quattro stagioni teatrali. Pur non essendo un periodo molto lungo, come richiesto dall'approccio di Bourdieu, è utile ad avere un'idea delle tendenze in atto. Infatti, le compagnie teatrali possono stabilire delle coproduzioni in uno o più stagioni, i *pattern* di connessioni possono cambiare nel tempo.

In quest'ottica, il cambiamento deve esser letto in termini "relazionali", considerando la dinamica della struttura del campo, osservando in che misura i leader e gli attori dominati mantengono i loro ruoli ricoperti nel corso del tempo.

Pertanto, un approccio longitudinale consente di esaminare in che misura questa struttura è stabile nel tempo, anche in termini di posizioni degli attori più/meno importanti. A tal fine si utilizza l'AFM (87, 163), che la gestione tabelle di dati a più vie, in questo caso, per rappresentare i diversi set di dati per ogni stagione teatrale.

Nello specifico l'applicazione dell'AFM per l'analisi dei network *two-mode* che variano nel tempo, (*two-mode time varying*), che includono come variabili supplementari anche le posizioni degli attori/eventi, consente di: (i) analizzare la struttura delle relazioni incorporata in ogni matrice di affiliazione, relativa ad ogni singola occasione temporale, in termini di similarità relazionale tra attori/eventi, (*analisi parziale*); (ii) analizzare in modo cumulativo le occasioni, cercando la

struttura relazionale complessiva racchiusa nella matrice di affiliazione generale, (considerando anche la collocazione delle posizioni degli eventi/attori in relazione agli attori/eventi a cui sono associati e rispetto alle loro caratteristiche descrittive) (*analisi globale*); (iii) evidenziare le differenze tra i vari punti temporali in termini di struttura globale (variazioni strutturali tra le occasioni, osservando i cambiamenti nel *pattern* relazionali sia nel tempo sia nei due modi); (iv) analizzare le variazioni degli attori/eventi nelle varie occasioni (osservando le variazioni nei *pattern* relazionali proiettando le matrici di affiliazione pesate di ogni occasione sul piano fattoriale globale). Gli ultimi due punti, (iii) e (iv), permettono di studiare il cambiamento longitudinale e il comportamento nel tempo dei modelli relazionali degli attori/eventi e delle posizioni a cui sono associati (169)¹.

5.4.2 Network *Two-mode* delle stagioni di co-produzioni

Il campo delle produzioni teatrali viene ricostruito attraverso il network di affiliazione delle istituzioni e delle compagnie che collaborano per mettere in scena le co-produzioni nella Regione Campania, in Italia (183). I dati sono costituiti da un set principale che contiene le informazioni sul network dei legami tra le compagnie e le co-produzioni a cui hanno partecipato, insieme con gli attributi descrittivi di entrambi. In aggiunta sono presenti una serie di informazioni qualitative sulla base di interviste con alcuni membri delle compagnie emerse come più rilevanti (183). I dati comprendono le compagnie e le co-produzioni relative a quattro stagioni teatrali: 2011-2012, 2012-2013, 2013-2014, 2014-2015. Il dataset complessivo è composto da 20 attori focali (con una propria produzione e una sede sul territorio campano) a partire dai quali sono state ricostruite le coproduzioni e le compagnie partner, e 80 organizzazioni non focali teatrali o non teatrali, per un totale di 100 compagnie e 157 co-coproduzioni². Nella matrice delle coproduzioni teatrali sono state aggiunte numerose informazioni relative alle caratteristiche

¹Per un breve dettaglio tecnico si rinvia all'appendice 5.5, per una trattazione dettagliata sul metodo dell'AFM per network *two-mode time-varying* si rinvia a Ragozini et al. (169).

²Inizialmente gli attori focali erano 33, tuttavia, 13 sono risultati completamente isolati, in quanto non coinvolti in coproduzioni nei periodi considerati

5.4 Analisi del campo attraverso l'Analisi Fattoriale Multipla e il Blockmodeling: il caso della produzione teatrale

degli attori, ossia le compagnie, e degli evento, ossia gli spettacoli in coproduzione. Gli attributi descrittivi degli attori riguardano una serie di caratteristiche quali: il tipo di organizzazione; il tipo di città (capoluogo di regione, capoluogo di provincia, e l'area provinciale); la dimensione e il tipo di sede del teatro principale (teatro storico, teatro moderno, costruzione non-teatrale riadattata per ospitare spettacoli dal vivo). Per quanto riguarda gli attributi descrittivi delle coproduzioni sono stati considerati: il genere teatrale (commedia, *cabaret*, contemporaneo, *musical*, teatro giovanile, sperimentale, ecc.); le caratteristiche dello script ¹ (originale, rivisitato, esistente non rivisitato); la presenza di elementi di multidisciplinarietà (combinazione di forme di arte, media e comunicazione); la presenza, o meno, di interazione con gli spettatori (183). Questi attributi vengono considerati come risorse la cui distribuzione definisce la struttura oggettiva dello spazio delle posizioni, pertanto vengono utilizzati come *proxy* per alcune delle forme di capitale indicate da Bourdieu (39). Nello specifico, per quanto riguarda le compagnie, il riconoscimento istituzionale come "Teatro Stabile" attiene al prestigio dell'organizzazione e costituisce una risorsa di capitale simbolico, gli altri tipi di organizzazioni teatrali, pur non essendo necessariamente meno prestigiose, mancano di questo riconoscimento. La dimensione della sede viene considerata come *proxy* per il capitale economico (non avendo a disposizione i dati sui finanziamenti), i teatri più grandi, in genere, possono contare su maggiori incassi, sia al botteghino sia per le sovvenzioni, rispetto a quelli più piccoli. In merito al capitale sociale, esso viene considerato nella sua forma inter-organizzativa e misurato come quantità di contatti con cui una compagnia può entrare in contatto tramite la condivisione di coproduzioni comuni e dai quali può trarre potenzialmente profitto. Per quanto riguarda gli attributi delle co-produzioni i generi e le caratteristiche delle *performance* esprimono il capitale simbolico che esse possono fornire alle compagnie. In particolare, le imprese teatrali ottenere il massimo profitto simbolico, contrariamente al profitto economico (35), dalle prestazioni più rari e per lo più non convenzionali o da opere teatrali che non hanno ancora ottenuto grande popolarità (183). Tutti questi attributi sono stati inseriti nell'analisi, rispettivamente, come variabili e punti supplementari (Paragrafo 3.4). Alla matrice di affiliazione finale, analizzata attraverso l'MFA, seguendo la

¹Lo script si riferisce l'opera letteraria o dramma su cui si basa una certa produzione

stessa logica, sono stati aggiunti, come variabili e punti supplementari, anche le partizioni derivanti dall'analisi di Blockmodeling che ha permesso di individuare i cluster equivalenti di attori e degli eventi ad essi associati, ossia le posizioni di network delle compagnie e i cluster delle coproduzioni.

5.4.3 *Two-mode* Blockmodeling delle stagioni di co-produzioni teatrali

Al fine di evidenziare le partizioni e le posizioni nella struttura del network di affiliazione, viene utilizzato il *two-mode* Blockmodeling (28, 81, 83)), che, come si è visto prevede una approccio diretto e consente di partizionare in modo diverso ma allo stesso tempo entrambi i modi del network. Si ricorda che nel *two-mode* Blockmodeling righe colonne hanno una natura diversa, pertanto, non si applica il concetto di blocchi diagonali per descrivere le relazione all'interno delle posizioni, ossia all'interno di un cluster. Nel *two-mode* Blockmodeling un blocco, ossia il cluster di relazioni è sempre definito da due cluster, uno corrispondente agli attori e l'altro agli eventi ad essi associati. Ogni blocco stabilito empiricamente riflette un modello di connessioni coerente con un certo tipo equivalenza.

Nel caso in esame i blocchi strutturalmente equivalenti corrispondono a modelli di connessioni in cui tutti i componenti del cluster delle compagnie sono legati a tutti i componenti del cluster delle coproduzioni, ossia sono tutti coinvolti nelle stesse coproduzioni. Lo stesso vale per le coproduzioni che devono coinvolgere le stesse compagnie. Blocchi regolarmente equivalenti corrispondono a modelli di connessioni in cui almeno un componente per ogni riga del cluster delle compagnie è legato con almeno un componente per ogni colonna del cluster delle coproduzioni, ossia ogni compagnia è coinvolta in almeno una coproduzione. Lo stesso vale per le coproduzioni che devono coinvolgere ognuna almeno una compagnia. Altri blocchi ideali sono stati coinvolti nell'analisi poichè quelli corrispondenti ai criteri di equivalenza strutturale e regolare non sono stati sufficienti a delineare i modelli di connessioni presenti nel network. Questi ultimi verranno discussi nel corso dell'analisi. Come si è ampiamente mostrato nel corso del secondo capitolo, infatti, nell'ambito del Blockmodeling generalizzato il ricorso ai tipi di blocchi ideali consente di definire nuove forme di equivalenza, che meglio

5.4 Analisi del campo attraverso l'Analisi Fattoriale Multipla e il Blockmodeling: il caso della produzione teatrale

riflettono i modelli di connessioni empirici, e che, nel suo insieme, solitamente, viene indicata come equivalenza generalizzata.

Nello studio della struttura del network delle co-produzioni teatrali si adotta l'approccio metodologico proposto nel capitolo precedente, in cui l'analisi di Blockmodeling è preceduto da una fase di analisi esplorativa attraverso l'AF. In questo modo è stabilito un Blockmodeling pre-specificato, in cui le informazioni empiriche e teoriche disponibili sul settore teatrale sono state integrate nella procedura di pre-specificazione al fine di definire un modello di struttura del network delle co-produzioni da "testare" per ogni stagione, in modo da valutare i cambiamenti della struttura relazionale nel tempo.

Più nel dettaglio, la pre-specificazione ha coinvolto alcune informazioni di carattere empiriche e teorico e si è articolata in tre tappe. Le informazioni hanno riguardato, da in lato, la conoscenza empirica ottenuta mediante l'analisi preliminare delle quattro stagioni di coproduzioni teatrali attraverso l'AFM e, al suo interno, le singole ACM relative ad ogni stagione. Dall'altro dalle conoscenze teoriche sul contesto della produzione culturale e delle caratteristiche del sistema teatrale regionale e nazionale (181, 182). I quattro blockmodel, uno per ogni stagione teatrale, sono stati stabiliti attraverso le seguenti tappe:

- (i) realizzazione dell'AFM che ha portato alla definizione di un numero di cluster di base per le compagnie e per le coproduzioni; all'individuazione di una serie di blocchi ideali possibili rispetto ai *pattern* di connessioni emergenti; alla definizione di un modello di struttura da utilizzare come base del blockmodel.
- (ii) definizione del modello di struttura con l'integrazione delle conoscenze teoriche, realizzazione dei blockmodel, ri-specificazione dei modelli di struttura secondo i tipi di blocchi ideali più appropriati ai modelli di connessioni.
- (iii) realizzazione un Blockmodeling pre-specificato con struttura del clustering parzialmente vincolata. Data l'esistenza di cluster delle compagnie con una composizione stabile nel tempo, sono stati imposti vincoli al clustering per far in modo che comparissero sempre nello stesso cluster, nello stesso ordine in modo da poter confrontare agevolmente le strutture nei quattro momenti

5.4 Analisi del campo attraverso l'Analisi Fattoriale Multipla e il Blockmodeling: il caso della produzione teatrale

	2011-2012C.	2012-2013C.	2013-2014C.	2014-2015C.
Compagnie	33	44	48	58
Co-produzioni	37	45	40	35
Nodi	70	89	88	93
Legami	84	109	95	97
Densità	0,07	0,06	0,05	0,05
Tasso medio di affiliazione	2,45	2,48	1,98	1,67
Ampiezza media degli eventi	2,27	2,42	2,38	2,77

	2011-2012S.	2012-2013S.	2013-2014S.	2014-2015S.
Compagnie	27	40	34	47
Co-produzioni	34	43	33	29
Nodi	61	83	67	76
Legami	78	105	79	85
Densità	0,08	0,06	0,07	0,06
Tasso medio di affiliazione	2,89	2,63	2,32	1,81
Ampiezza media degli eventi	2,29	2,44	2,39	2,93

Tabella 5.1: Statistiche descrittive dei network e dei sub-network delle co-produzioni teatrali per le quattro stagioni. C. indica l'intero network, S. indica il sub-network.

temporali e poter identificare chiaramente (attraverso le etichette) i cluster da inserire nell'AFM finale.

Delle tappe elencate, di seguito vengono presentati esclusivamente i risultati relativi ai blockmodel finali.

Nella Figura 5.1 vengono riportate alcune statistiche descrittive per i network completi e per i sub-network delle coproduzioni teatrali per le quattro stagioni. Come si può notare, tutti i network delle coproduzioni sono caratterizzati da valori molto bassi sia per la densità sia per il tasso medio di partecipazione e per l'ampiezza media degli eventi. Com'è noto, l'analisi di Blockmodeling risulta complessa per le reti caratterizzate da una bassa densità, con legami sparsi. Pertanto, per agevolare l'operazione di partizionamento, si è scelto di condurre l'analisi su un sub-network estratto attraverso la tecnica di decomposizione in *weak component*, selezionandone il numero maggiore possibile in modo da mantenere una struttura vicina a quella iniziale. Tutti i modelli che seguono sono stati elaborati attraverso la procedura di *two-mode* Blockmodeling generalizzato implementata in Pajek (11, 16), con 1000 ripetizioni.

5.4 Analisi del campo attraverso l'Analisi Fattoriale Multipla e il Blockmodeling: il caso della produzione teatrale

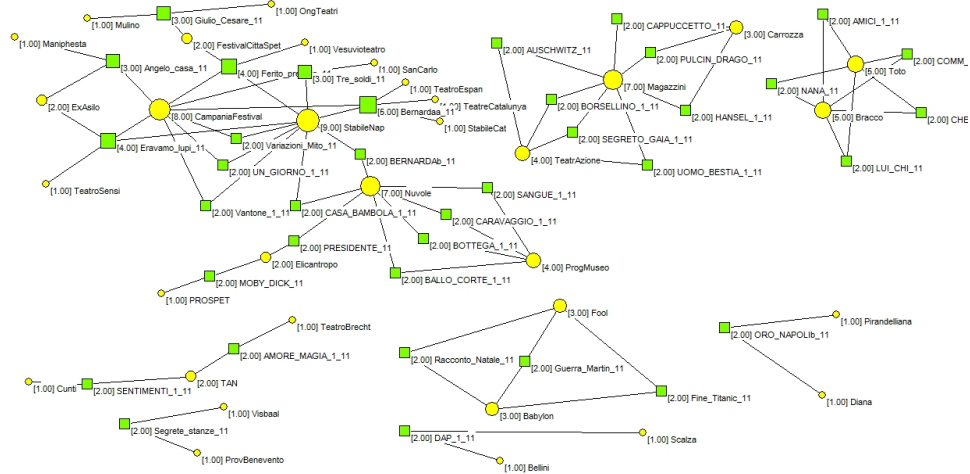


Figura 5.1: *Two-mode* network delle co-produzioni teatrali 2011-2012. *Degree centrality*. Cerchi: compagnie; quadrati: coproduzioni.

Blockmodeling della stagione di co-produzioni teatrali 2011-2012. Nella Figura 5.1 viene riportato il network *two-mode* delle co-produzioni della stagione in esame sulla base del *degree* delle compagnie e delle coproduzioni. Il network è caratterizzato da una certa frammentarietà con una componente principale e altre più piccole. Nella prima, Stabile Napoli, Campania Festival e Nuvole hanno un ruolo centrale in termini di numero di partecipazioni a coproduzioni e Bernarda.a, Ferito.preludio e Eravamo.lupi in termini di numero di partecipanti alla coproduzione. Nelle componenti più piccole emergono da un lato, Magazzini e, dall'altro Bracco e Totò. A partire da queste informazioni elementari è possibile immaginare un ruolo preminente di tali compagnie che, attraverso la condivisione di numerose produzioni, dominano gran parte del sistema di coproduzioni teatrali della stagione.

Il sub-network delle coproduzioni teatrali nella stagione 2011-2012 è caratterizzato da 27 organizzazioni e 34 spettacoli in coproduzione. Nel network sono presenti 78 legami, una densità dello 0,08, con un tasso medio di partecipazione degli attori (*affiliation average rate*) pari a 2,89, e un numero medio di attori per evento (*event average sizes*) pari a 2,29.

La Figura 5.2 riporta il modello di struttura definitivo del *two-mode* blockmodel costruito a partire dalle informazioni ottenute attraverso l'analisi preliminare

5.4 Analisi del campo attraverso l'Analisi Fattoriale Multipla e il Blockmodeling: il caso della produzione teatrale

tramite AFM, e nello specifico a partire dall'analisi delle coordinate parziali della stagione 2011-2012. Il modello ideale di blockmodel pre-specificato presenta una partizione in 6 cluster sia per le compagnie sia per le coproduzioni ed è caratterizzato da una struttura segmentata con blocchi disposti quasi esclusivamente lungo la diagonale, ad eccezione del blocco (6-3). Nella struttura del blockmodel è stata indicata la collocazione di tutti i tipi di blocchi e per ognuno di essi è stato specificato almeno un tipo di blocco ideale. I tipi di blocchi ideali selezionati, risultati più appropriati ai modelli di connessioni esistenti, sono principalmente: completo [comp], che corrisponde al criterio di equivalenza strutturale, in cui tutti i membri del cluster delle compagnie sono strutturalmente equivalenti in quanto condividono le stesse coproduzioni, e viceversa per le coproduzioni; denso [den], che corrisponde ad una sorta di grado di equivalenza strutturale, in cui i membri del cluster delle compagnie condividono la maggior parte delle coproduzioni (ossia vi è una densità di legami $\gamma > 0,75$), e viceversa per le coproduzioni; *col-regular* [cre], che corrisponde ad un rilassamento del criterio dell'equivalenza regolare, in cui ogni componente del cluster delle coproduzioni è coinvolto da almeno un membro del cluster delle compagnie; nullo [-], in cui non sono presenti legami tra compagnie e coproduzioni.

Attraverso un modello ideale di blockmodel di questo tipo si ipotizza una struttura del sistema di coproduzioni teatrali fortemente segmentata, con alcune aree caratterizzate da elevata collaborazione, alla quale potrebbe corrispondere una differenziazione per generi teatrali, per distribuzione delle risorse, per caratteristiche delle compagnie, ecc. Queste caratterizzazioni, in questa fase, non possono essere ottenute direttamente tramite la procedura di Blockmodeling e verranno esplorate in seguito nell'AFM finale in cui le posizioni e i cluster estratti verranno messi in relazione alla struttura oggettiva definita dalla distribuzione delle varie forme di risorse presenti nel campo.

Nella Figura 5.3 vengono riportati i risultati del Blockmodeling (6-6)-*partition*. Come si può notare dalle immagini (a) e (b) la procedura partiziona il network in 2 blocchi densi, 3 completi e 1 *col-regular*, con un valore per la funzione criterio pari a 13 inconsistenze. Il blockmodel empirico sostanzialmente conferma la struttura del modello teorico, ad eccezione del blocco definito dai cluster (6-3) individuato come blocco nullo. Nel blockmodel finale il primo blocco completo

5.4 Analisi del campo attraverso l'Analisi Fattoriale Multipla e il Blockmodeling: il caso della produzione teatrale

Image matrix definition:

	1	2	3	4	5	6
1:	[com den]	[-]	[-]	[-]	[-]	[-]
2:	[-]	[com]	[-]	[-]	[-]	[-]
3:	[-]	[-]	[com den]	[-]	[-]	[-]
4:	[-]	[-]	[-]	[com den]	[-]	[-]
5:	[-]	[-]	[-]	[-]	[com]	[-]
6:	[-]	[-]	[- cre]	[-]	[-]	[- cre]

Figura 5.2: Modello di struttura del blockmodel delle co-produzioni teatrali 2011-2012.

è definito dai cluster (2-2) e comprende le compagnie Bracco e Totò; il secondo blocco completo è definito dai cluster (3-3) e coinvolge le compagnie Campania Festival e Stabile Napoli; il terzo blocco completo è definito dai cluster (5-5) e include le compagnie Babylon e Fool. Ognuna delle compagnie presenti in ciascuno dei cluster che definiscono ciascun blocco condivide le stesse coproduzioni e sono caratterizzate da una relazione di equivalenza strutturale esatta. Per quanto riguarda i due blocchi densi, il primo è definito dai cluster (1-1) e comprende le compagnie ProgMuseo e Nuvole; il secondo è definito dai cluster (4-4) e coinvolge le compagnie Teatro Azione e Magazzini. Le compagnie appartenenti ai cluster che determinano tali blocchi sono caratterizzati da una intensa collaborazione nella maggior parte delle coproduzioni incluse nei cluster a cui sono associate. Infine, l'unico blocco *col-regular* è definito dai cluster (6-6) ed è caratterizzato da un modello di connessioni con collaborazioni sparse, in cui, ogni coproduzione è coinvolta almeno da una compagnia del cluster di attori corrispondente. Questo blocco evidenzia un modello di connessioni frammentato con una scarsa intensità di collaborazione. In definitiva, il blockmodel empirico mostra una struttura simile a quella per sottogruppi coesi, caratterizzata da diversi attori *core* che attraggono la maggior parte delle coproduzioni e assumono una posizione di rilievo nell'intero sistema. Tuttavia, questa propensione alla collaborazione è limitata ai componenti dei singoli sottogruppi e non coinvolge (o coinvolge troppo poco per essere considerata come relazione strutturale) attori esterni appartenenti a cluster diversi, pertanto non si traduce in fattore di integrazione del sistema ma favorisce una struttura differenziata. In ultimo, nell'immagine (c) della Figura 5.3, è interessante osservare la localizzazione di alcune delle 13 inconsistenze rilevate dalla funzione criterio. Due di esse sono collocate nel blocco definito dai cluster

5.4 Analisi del campo attraverso l'Analisi Fattoriale Multipla e il Blockmodeling: il caso della produzione teatrale

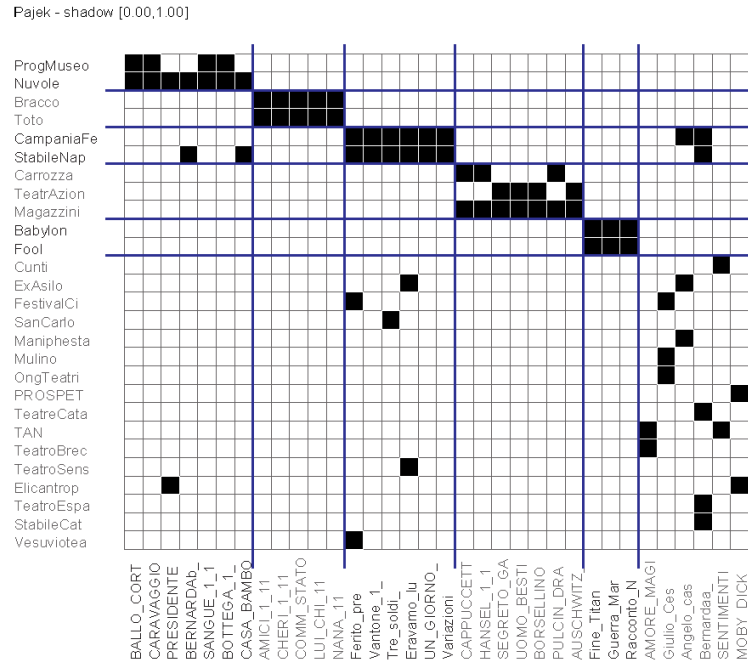
(3-1) e sono date dai legami di collaborazione tra due compagnie rilevanti Nuvole e Stabile Napoli; 8 sono collocate tra i blocchi definiti dai cluster (6-3) e (3-6) e sono date dai legami di collaborazione tra Campania Festival e Stabile Napoli con il gruppo delle compagnie più "periferiche" all'interno del sistema. Queste connessioni pur non essendo state identificate dal blockmodel come coerenti con un tipo di blocco specifico, suggeriscono una possibile tendenza ad una maggiore apertura alla collaborazione esterna, la cui esistenza sarà valutata nei blockmodel delle stagioni successive.

Blockmodeling della stagione di co-produzioni teatrali 2012-2013. Nella Figura 5.4 viene riportato il network *two-mode* delle co-produzioni della stagione 2012-2013 sulla base del *degree* delle compagnie e delle coproduzioni. Il network è caratterizzato da un grado di connessione maggiore rispetto alla stagione precedente, con una componente principale che comprende la quasi totalità dei nodi e altre tre più piccole. Nella prima, Stabile Napoli, Nuvole, Campania Festival e Magazzini hanno un ruolo centrale in termini di numero di partecipazioni a coproduzioni e Fabula, Contratto, Ferito, Bisbetica e Biodiversity in termini di numero di partecipanti alla coproduzione. Nelle componenti più piccole emergono Bracco e Totò. Si conferma il ruolo preminente delle compagnie individuate nella precedente stagione, alle quali si unisce Magazzini. Queste compagnie attraverso la condivisione di una serie di produzioni creano una catena di connessioni che favorisce una maggiore integrazione del sistema.

Il sub-network delle coproduzioni teatrali nella stagione 2011-2012 è caratterizzato da 40 organizzazioni e 43 spettacoli in coproduzione. Nel network sono presenti 105 legami, una densità dello 0,06, con un tasso medio di partecipazione degli attori (*affiliation average rate*) pari a 2,63, e un numero medio di attori per evento (*event average sizes*) pari a 2,44.

La Figura 5.5 riporta il modello di struttura definitivo del *two-mode* blockmodel costruito a partire dalle informazioni ottenute attraverso l'analisi preliminare tramite AFM, e nello specifico a partire dall'analisi delle coordinate parziali della stagione 2012-2013. Il modello ideale di blockmodel pre-specificato presenta una partizione in 6 cluster per le compagnie e 5 per le coproduzioni ed è caratterizzato da una struttura che combina quella segmentata con quella del modello

5.4 Analisi del campo attraverso l'Analisi Fattoriale Multipla e il Blockmodeling: il caso della produzione teatrale



(a) Matrice delle co-produzioni teatrali 2011-2012 con partizioni *two-mode* blockmodel

Final Image Matrix:

	1	2	3	4	5	6
1	den	-	-	-	-	-
2	-	com	-	-	-	-
3	-	-	com	-	-	-
4	-	-	-	den	-	-
5	-	-	-	-	com	-
6	-	-	-	-	-	cre

(b) Matrice immagine

Final Error Matrix:

	1	2	3	4	5	6
1	0	0	0	0	0	0
2	0	0	0	0	0	0
3	2	0	0	0	0	3
4	0	0	0	2	0	0
5	0	0	0	0	0	0
6	1	0	5	0	0	0

Final error = 13.000

(c) Inconsistenze e Funzione criterio

Figura 5.3: *Two-mode* Blockmodeling generalizzato delle co-produzioni teatrali 2011-2012.

core-periphery. In essa i blocchi sono disposti lungo la diagonale, ad eccezione di quello (6-3). Come per il Blockmodeling della precedente stagione, nella struttura del blockmodel è stata indicata la collocazione di tutti i tipi di blocchi e per ognuno di essi è stato specificato almeno un tipo di blocco ideale. I tipi di blocchi ideali selezionati, risultati più appropriati ai modelli di connessioni esistenti, sono: completo [comp], denso [den], *col-regular* [cre], nullo [-], e, in aggiunta, *row-regular* [rre] che corrisponde ad un rilassamento del criterio dell'equivalenza

5.4 Analisi del campo attraverso l'Analisi Fattoriale Multipla e il Blockmodeling: il caso della produzione teatrale

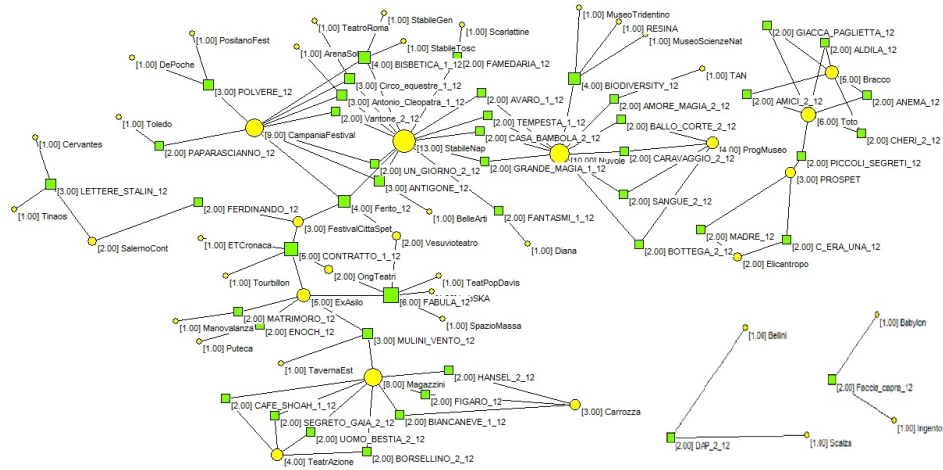


Figura 5.4: *Two-mode* network delle co-produzioni teatrali 2012-2013. *Degree centrality*. Cerchi: compagnie; quadrati: coproduzioni.

Image matrix definition:

	1	2	3	4	5
1: [com den]	[-]	[-]	[-]	[-]	[-]
2: [-]	[com]	[-]	[-]	[-]	[-]
3: [-]	[-]	[com den]	[-]	[-]	[-]
4: [-]	[-]	[-]	[com den]	[-]	[-]
5: [-]	[-]	[-]	[-]	[com cre]	[-]
6: [-]	[-]	[- rre]	[-]	[-]	[-]

Figura 5.5: Modello di struttura del blockmodel delle co-produzioni teatrali 2012-2013.

regolare, in cui ogni componente del cluster delle compagnie è coinvolto in almeno una produzione del cluster delle coproduzioni.

Attraverso un modello ideale di blockmodel di questo tipo si ipotizza una struttura del sistema di coproduzioni teatrali ancora una volta segmentata, ma con l'aggiunta di spazi di collaborazione tra cluster di compagnie diverse sulle stesse coproduzioni, che sottendono una parallela gerarchizzazione della struttura.

Nella Figura 5.6 vengono riportati i risultati del Blockmodeling (6-5)-*partition*. Come si può notare dalle immagini (a) e (b) la procedura partiziona le connessioni presenti nel network in 1 blocco completo, 3 densi, 1 *col-regular* e 1 *row-regular*, con un valore per la funzione criterio pari a 12 inconsistenze. Il blockmodel empirico conferma la struttura del modello teorico. Nel blockmodel finale l'unico blocco completo è definito dai cluster (2-2) e comprende le compagnie Bracco e

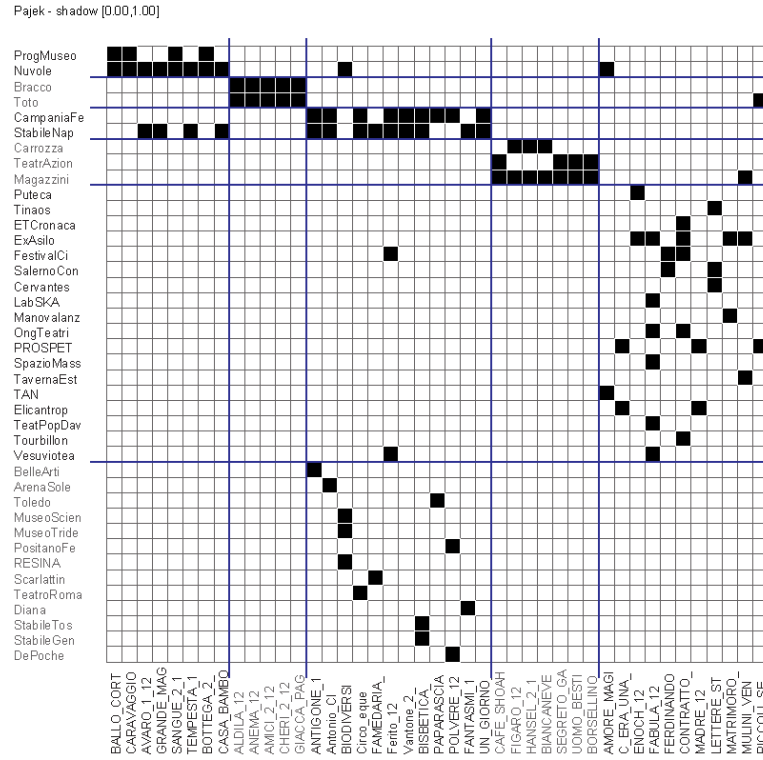
5.4 Analisi del campo attraverso l'Analisi Fattoriale Multipla e il Blockmodeling: il caso della produzione teatrale

Totò che sono caratterizzati da una relazione di equivalenza strutturale esatta. Per quanto riguarda i blocchi densi, il primo è definito dai cluster (1-1) e comprende le compagnie ProgMuseo e Nuvole; il secondo è definito dai cluster (3-3) e coinvolge le compagnie Campania Festival e Stabile Napoli; il terzo è definito dai cluster (4-4) e comprende le compagnie Teatro Azione, Magazzini e Carrozza. Le compagnie appartenenti ai cluster che determinano tali blocchi sono caratterizzati da una intensa collaborazione nella maggior parte delle coproduzioni incluse nei cluster a cui sono associate. Il blocco *col-regular* è definito dai cluster (5-5) ed è caratterizzato da un modello di connessione con collaborazioni sparse, in cui, ogni coproduzione è coinvolta almeno da una compagnia del cluster di organizzazioni corrispondente. Infine, e a differenza della stagione precedente, il blocco definito dai cluster (6-3) viene identificato come consistente con il modello di connessioni alla base del tipo di blocco ideale *row-regular*. A differenza del blocco (5-5) che sembra definire un'area di collaborazione periferica (in termini di intensità di collaborazione, eccetto alcune compagnie come Ex-Asilo e Prospet) rispetto all'intero sistema di coproduzioni, con una sua relativa autonomia, il blocco (6-3) individua uno spazio di collaborazione periferico rispetto ad un *core* specifico, costituito da Campania Festival e Stabile Napoli. Questo tipo di relazione strutturale rimanda a dinamiche di natura diversa, gerarchiche, a cui sono associati sia vincoli sia opportunità.

In definitiva, il blockmodel empirico conferma una struttura simile a quella per sottogruppi coesi, caratterizzata da diversi attori *core* che attraggono un gran numero di coproduzioni e continuano ad assumere una posizione di rilievo nel sistema di coproduzioni anche nella stagione 2012-2013. Inoltre, a differenza della stagione precedente emerge una maggiore propensione alla collaborazione con attori esterni appartenenti a cluster diversi, relazioni che assumono un carattere strutturale. Tale tendenza favorisce una maggiore integrazione del sistema di coproduzioni. Infine, nella matrice delle degli errori (c) della Figura 5.6, è interessante osservare la localizzazione di 4 delle 13 inconsistenze rilevate dalla funzione criterio. Esse sono collocate nel blocco definito dai cluster (3-1) e sono date dai legami di collaborazione tra Nuvole e Stabile Napoli, di cui 2 già presenti nella stagione precedente. Tale relazione anche se non assume una rilevanza

5.4 Analisi del campo attraverso l'Analisi Fattoriale Multipla e il Blockmodeling: il caso della produzione teatrale

strutturale riveste una certa importanza in termini di connessioni tra differenti aree dominanti del sistema.



(a) Matrice delle co-produzioni teatrali 2012-2013 con partizioni *two-mode* blockmodel

Final Image Matrix:

	1	2	3	4	5
1	den	-	-	-	-
2	-	com	-	-	-
3	-	-	den	-	-
4	-	-	-	den	-
5	-	-	-	-	cre
6	-	-	rre	-	-

(b) Matrice immagine

Final Error Matrix:

	1	2	3	4	5
1	0	0	1	0	1
2	0	0	0	0	1
3	4	0	0	0	0
4	0	0	0	2	1
5	0	0	2	0	0
6	0	0	0	0	0

Final error = 12.000

(c) Inconsistenze e Funzione criterio

Figura 5.6: *Two-mode* Blockmodeling generalizzato delle co-produzioni teatrali 2012-2013.

5.4 Analisi del campo attraverso l'Analisi Fattoriale Multipla e il Blockmodeling: il caso della produzione teatrale

Blockmodeling della stagione di co-produzioni teatrali 2013-2014. Nella Figura 5.7 viene riportato il network *two-mode* delle co-produzioni della stagione 2013-2014 sulla base del *degree* delle compagnie e delle coproduzioni. Il network mostra una certa variazione rispetto alla stagione precedente. Essa presenta sempre una componente principale che comprende la maggior parte dei nodi e altre più piccole. Tuttavia, emerge una riduzione complessiva dei *degree* sia per le compagnie che per le coproduzioni. Nella componente principale, Nuvole, Stabile Napoli e Magazzini hanno una certa importanza in termini di numero di partecipazioni a coproduzioni, ad esse si aggiungono anche Campani Festival e Totò la cui componente si connette a quella principale. Inoltre, Macaluso, Bisbetica, Titatic.neiwiller, Biodiversity, Barbone e Nessuno assumono un ruolo importante in termini di numero di partecipanti alla coproduzione. Anche in questo caso, emerge il ruolo preminente delle compagnie individuate nella precedente stagione, con una riduzione dell'intensità di collaborazione per Campani Festival.

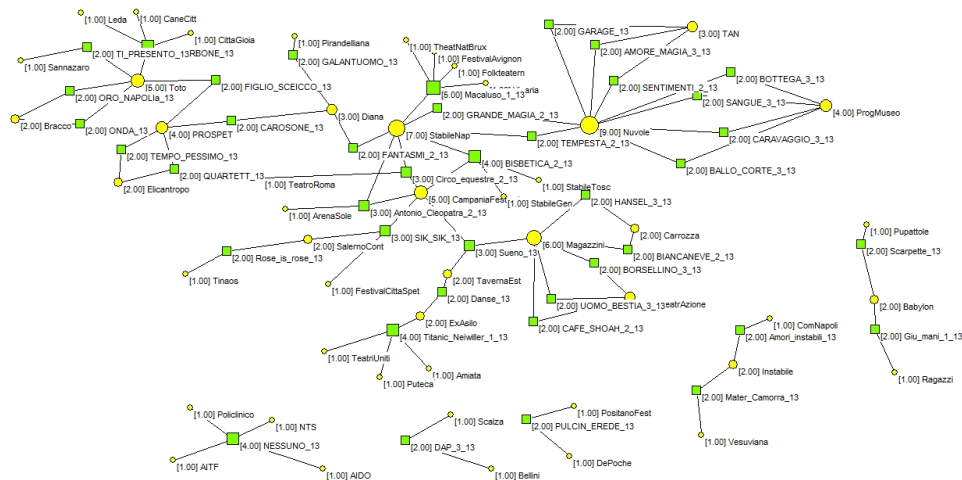


Figura 5.7: *Two-mode* network delle co-produzioni teatrali 2013-2014. *Degree centrality*. Cerchi: compagnie; quadrati: coproduzioni.

Il sub-network delle coproduzioni teatrali nella stagione 2013-2014 è caratterizzato da 34 organizzazioni e 33 spettacoli in coproduzione. Nel network sono presenti 79 legami, una densità dello 0,07, con un tasso medio di partecipazione degli attori (*affiliation average rate*) pari a 2,32, e un numero medio di attori per evento (*event average sizes*) pari a 2,39.

5.4 Analisi del campo attraverso l'Analisi Fattoriale Multipla e il Blockmodeling: il caso della produzione teatrale

Image matrix definition:

	1	2	3	4	5
1:	[com den]	[-]	[-]	[-]	[-]
2:	[-]	[com den]	[-]	[-]	[-]
3:	[-]	[-]	[com den]	[-]	[-]
4:	[-]	[-]	[-]	[com den]	[-]
5:	[-]	[-]	[- cre]	[-]	[- cre den]

Figura 5.8: Modello di struttura del blockmodel delle co-produzioni teatrali 2013-2014.

La Figura 5.8 riporta il modello di struttura definitivo del *two-mode* blockmodel costruito a partire dalle informazioni ottenute attraverso l'analisi preliminare tramite AFM, e nello specifico a partire dall'analisi delle coordinate parziali della stagione 2013-2014. Il modello ideale di blockmodel pre-specificato presenta una partizione in 5 cluster sia per le compagnie sia per le coproduzioni e, come nella stagione precedente, è caratterizzata da una struttura che combina un modello segmentato, simile a quello dei sottogruppi coesi, con quello gerarchico, in linea con quello del *core-periphery*. In essa i blocchi sono disposti lungo la diagonale, ad eccezione di quello (5-3). Come per il Blockmodeling della precedente stagione, nella struttura del blockmodel è stata indicata la collocazione di tutti i tipi di blocchi e per ognuno di essi è stato specificato almeno un tipo di blocco ideale. I tipi di blocchi ideali selezionati sono: completo [comp], denso [den], *col-regular* [cre] e nullo [-]. Anche in questo caso, con un modello ideale di blockmodel di questo tipo si ipotizza una struttura del sistema di coproduzioni teatrali che combina la differenziazione con la collaborazione tra cluster di compagnie diverse nelle stesse coproduzioni.

Nella Figura 5.9 vengono riportati i risultati del Blockmodeling (5-5)-*partition*. Come emerge dalle immagini (a) e (b) la procedura partiziona le connessioni presenti nel network in 4 blocchi densi e 2 *col-regular* con un valore per la funzione criterio pari a 12 inconsistenze e 6 soluzioni ugualmente ben modellate ad indicare modelli di connessioni meno definiti. Il blockmodel empirico conferma la struttura del modello teorico. In particolare, nel blockmodel finale, dei quattro blocchi densi il primo è definito dai cluster (1-1) e comprende le compagnie ProgMuseo e Nuvole; il secondo è definito dai cluster (2-2) e comprende le compagnie Bracco e Totò; il terzo è definito dai cluster (3-3) e coinvolge le compagnie Campania Festival e Stabile Napoli; il quarto è definito dai cluster (4-4) e comprende le

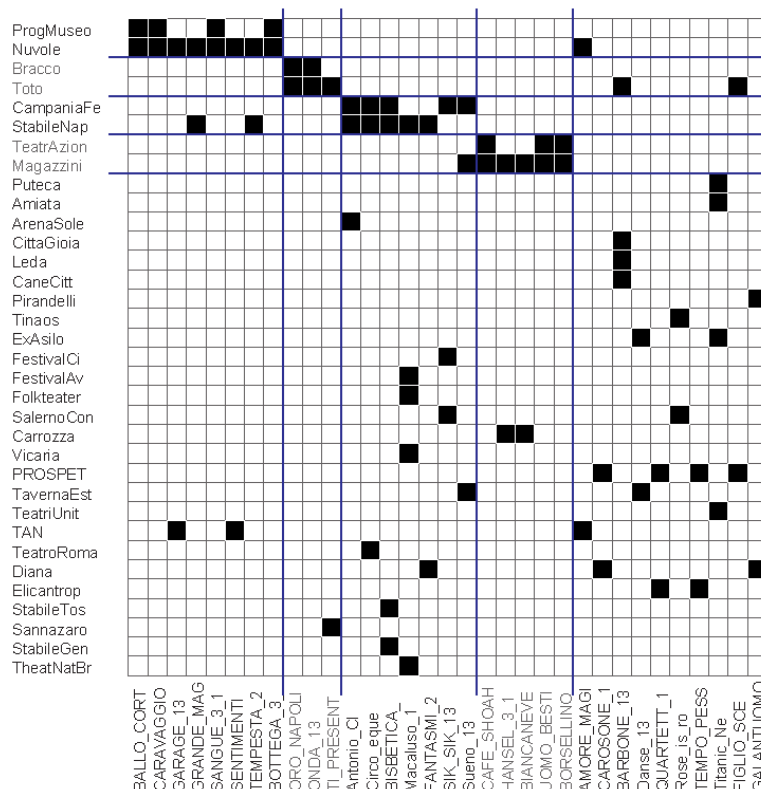
5.4 Analisi del campo attraverso l'Analisi Fattoriale Multipla e il Blockmodeling: il caso della produzione teatrale

compagnie Teatro Azione e Magazzini. Le compagnie appartenenti ai cluster che determinano tali blocchi sono caratterizzati da una intensa collaborazione nella maggior parte delle coproduzioni incluse nei cluster a cui sono associate. I due blocchi *col-regular* sono definiti dai cluster (5-5) e (5-3) e sono caratterizzati da un modello di connessione con collaborazioni sparse, in cui, ogni coproduzione è coinvolta almeno da una compagnia del cluster di organizzazioni corrispondente. Anche in questo caso il blocco (5-3) individua uno spazio di collaborazione periferico rispetto al *core* costituito da Campania Festival e Stabile Napoli. Di questo cluster 4 legami erano già presenti nella stagione precedente e più che periferici essi costituiscono un'apertura verso l'esterno dei confini regionali, infatti definiscono legami di collaborazione con altre organizzazioni di una certa rilevanza nazionale, esterne al sistema teatrale campano (TeatroRoma, StabileGen, StabileTos). In definitiva, il blockmodel empirico conferma una struttura mista con sottogruppi coesi, che collaborano con una certa intensità alle stesse coproduzioni, e *core-periphery* con collaborazioni con attori esterni appartenenti a cluster diversi, relazioni che assumono un carattere strutturale. Infine, per quanto riguarda la localizzazione di alcune delle inconsistenze (immagine (c) Figura 5.9), è possibile osservare che 2 di esse sono collocate nel blocco definito dai cluster (3-1) e sono date dai legami di collaborazione tra Nuvole e Stabile Napoli, altre 2 di esse sono collocate nel blocco definito dai cluster (2-5) e sono date dai legami di collaborazione tra Totò e il cluster delle compagnie con collaborazioni sparse.

Blockmodeling della stagione di co-produzioni teatrali 2014-2015. Nella Figura 5.10 viene riportato il network *two-mode* delle co-produzioni della stagione 2014-2015 sulla base del *degree* delle compagnie e delle coproduzioni. Il network mostra una variazione accentuata rispetto alle stagioni precedenti. Essa presenta ugualmente una componente principale che include gran parte dei nodi, e allo stesso tempo, rispetto alla stagione precedente, presente un numero maggiore di piccole componenti. Dal network emerge la scomparsa di Bracco e Totò e un ridimensionamento dell'importanza di Campania Festival. Nella componente principale, Stabile Napoli e Nuvole hanno una certa importanza in termini di numero di partecipazioni a coproduzioni e ad esse si aggiungono anche Magazzini, Ex-Asilo e ProgMuseo. Inoltre, Frankenstein, Macaluso, Panni degli altri, Velo e

5.4 Analisi del campo attraverso l'Analisi Fattoriale Multipla e il Blockmodeling: il caso della produzione teatrale

Pajek - shadow [0.00,1.00]



(a) Matrice delle co-produzioni teatrali 2013-2014 con partizioni *two-mode* blockmodel

Final Image Matrix
(for the first obtained solution):

	1	2	3	4	5
1	den	-	-	-	-
2	-	den	-	-	-
3	-	-	den	-	-
4	-	-	-	den	-
5	-	-	cre	-	cre

(b) Matrice immagine

Final Error Matrix
(for the first obtained solution):

	1	2	3	4	5
1	0	0	0	0	1
2	0	0	0	0	2
3	2	0	1	0	0
4	0	0	1	0	0
5	2	1	0	2	0

Final error = 12.000
(6 solutions)

(c) Inconsistenze e Funzione criterio

Figura 5.9: *Two-mode* Blockmodeling generalizzato delle co-produzioni teatrali 2013-2014.

Contratto assumono un ruolo importante in termini di numero di partecipanti alla coproduzione. Tali variazioni potrebbero segnalare una tendenza al cambiamento in atto verso una riconfigurazione dei rapporti collaborazione.

5.4 Analisi del campo attraverso l'Analisi Fattoriale Multipla e il Blockmodeling: il caso della produzione teatrale

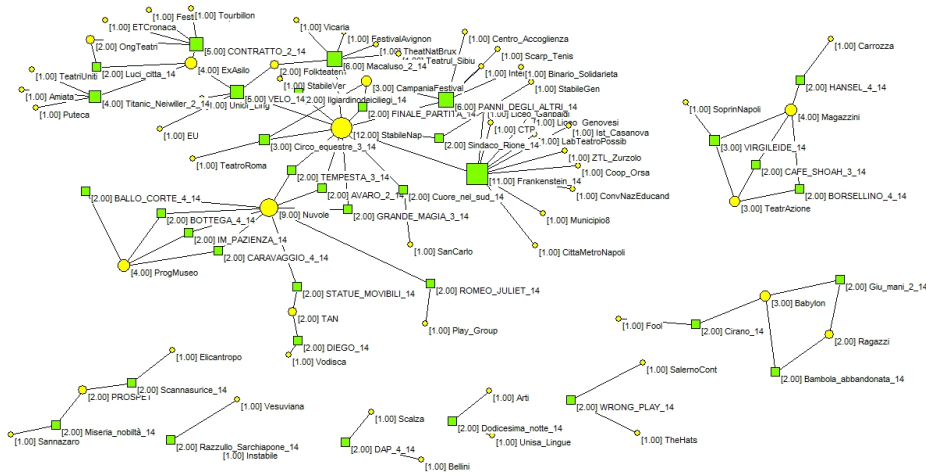


Figura 5.10: *Two-mode* network delle co-produzioni teatrali 2014-2015. *Degree centrality*. Cerchi: compagnie; quadrati: coproduzioni.

Il sub-network delle coproduzioni teatrali nella stagione 2014-2015 è caratterizzato da 47 organizzazioni e 29 spettacoli in coproduzione. Nel network sono presenti 85 legami, una densità dello 0,06, con un tasso medio di partecipazione degli attori (*affiliation average rate*) pari a 1,81, e un numero medio di attori per evento (*event average sizes*) pari a 2,93.

La Figura 5.11 riporta il modello di struttura definitivo del *two-mode* blockmodel costruito a partire dalle informazioni ottenute attraverso l'analisi preliminare tramite AFM, e nello specifico a partire dall'analisi delle coordinate parziali della stagione 2014-2015. Il modello ideale di blockmodel pre-specificato presenta una partizione in 6 cluster per le compagnie e 5 per le coproduzioni ed è caratterizzata da una struttura di base simile a quella delle stagioni precedenti che combina un modello segmentato (gruppi coesi) con quello gerarchico (*core-periphery*). In essa i blocchi sono disposti lungo la diagonale e al di sotto di essa. Come per il Blockmodeling della precedente stagione, nella struttura del blockmodel è stata indicata la collocazione di tutti i tipi di blocchi e per ognuno di essi è stato specificato almeno un tipo di blocco ideale. I tipi di blocchi ideali selezionati sono: completo [comp], denso [den], nullo [-], *col-regular* [cre], *row-regular* [rre] e *row-functional* [rfn] che definisce un modello di connessioni in cui ogni compagnia di un cluster è coinvolto esattamente in un'unica coproduzione del cluster

5.4 Analisi del campo attraverso l'Analisi Fattoriale Multipla e il Blockmodeling: il caso della produzione teatrale

Image matrix definition:

	1	2	3	4	5
1:	[com den]	[-]	[-]	[-]	[-]
2:	[-]	[com den]	[-]	[-]	[-]
3:	[- rre]	[-]	[com den]	[-]	[-]
4:	[-]	[-]	[-]	[com den]	[-]
5:	[-]	[-]	[- rfn]	[-]	[com]
6:	[-]	[-]	[- cre]	[-]	[- cre]

Figura 5.11: Modello di struttura del blockmodel delle co-produzioni teatrali 2014-2015.

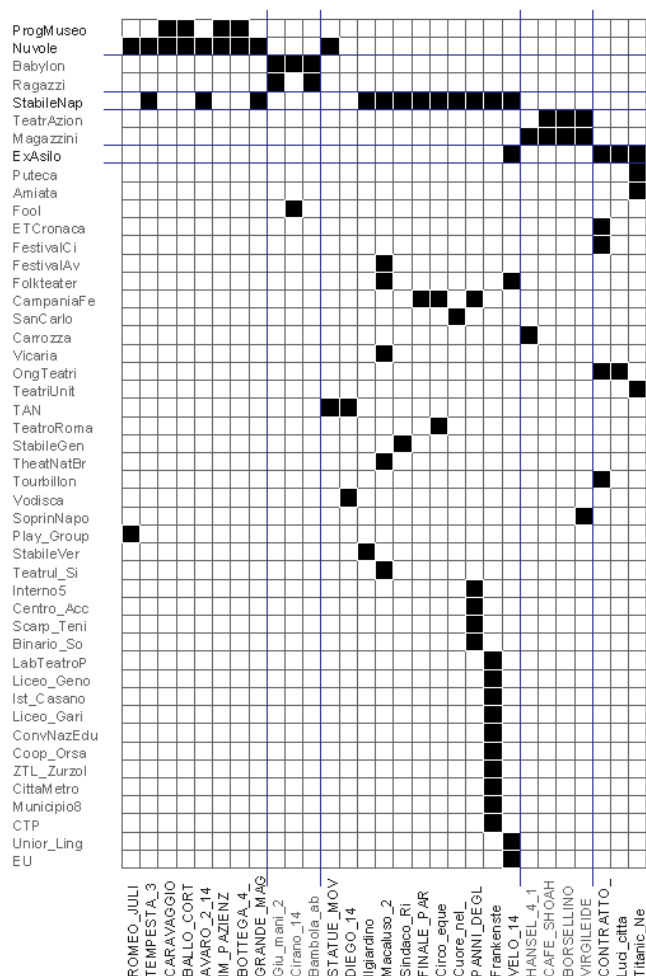
corrispondente. Anche in questo caso il modello ideale di blockmodel sottende una struttura del sistema di coproduzioni teatrali che combina la differenziazione con la collaborazione tra cluster di compagnie diverse nelle stesse coproduzioni.

Nella Figura 5.12 vengono riportati i risultati del Blockmodeling (6-5)-*partition*. Come emerge dalle immagini (a) e (b) la procedura partiziona le connessioni presenti nel network in un unico blocco completo, 4 blocchi densi, 2 *col-regular*, 1 *row-regular* e 1 *row-functional* con un valore per la funzione criterio pari a 5 inconsistenze e 2 soluzioni ugualmente ben modellate. Il blockmodel empirico conferma la struttura del modello teorico. In particolare, nel blockmodel finale, l'unico blocco completo è definito dai cluster (5-5) e comprende solo Ex-Asilo. Dei quattro blocchi densi il primo è definito dai cluster (1-1) e comprende le compagnie ProgMuseo e Nuvole; il secondo è definito dai cluster (2-2) e comprende le compagnie Babylon e Ragazzi; il terzo è definito dai cluster (3-3) e coinvolge l'organizzazione Stabile Napoli; il quarto è definito dai cluster (4-4) e comprende le compagnie Teatro Azione e Magazzini. Le compagnie appartenenti ai cluster che determinano tali blocchi sono caratterizzati da una intensa collaborazione nella maggior parte delle coproduzioni incluse nei cluster a cui sono associate. I due blocchi *col-regular* sono definiti dai cluster (6-3) e (6-5) e sono caratterizzati da un modello di connessione con collaborazioni sparse, in cui, ogni coproduzione è coinvolta almeno da una compagnia del cluster di organizzazioni corrispondente. In questo caso entrambi i blocchi individuano uno spazio di collaborazione periferico rispetto ad un *core*. Nel primo caso, (6-3), il centro è dato da Stabile Napoli e, per il secondo, (6-5), è dato da ExAsilo. Per quanto riguarda il blocco *row-regular*, esso è definito dai cluster (3-1) e coinvolge le collaborazioni di Stabile Napoli in alcune delle coproduzioni associate al cluster di ProgMuseo e Nuvole.

5.4 Analisi del campo attraverso l'Analisi Fattoriale Multipla e il Blockmodeling: il caso della produzione teatrale

Infine, il blocco *row-functional* è definito dai cluster (5-3) e comprende la collaborazione di ExAsilo in una delle coproduzioni associate al cluster di Stabile Napoli. Non si riscontrano peculiarità nella collocazione delle inconsistenze (immagine (c) Figura 5.12)

Pajek - shadow[0.00,1.00]



(a) Matrice delle co-produzioni teatrali 2014-2015 con partizioni *two-mode* blockmodel

Final Image Matrix
(for the first obtained solution):

	1	2	3	4	5
1	den	-	-	-	-
2	-	den	-	-	-
3	rre	-	den	-	-
4	-	-	-	den	-
5	-	-	rfrn	-	com
6	-	-	cre	-	cre

Final Error Matrix
(for the first obtained solution):

	1	2	3	4	5
1	0	0	1	0	0
2	0	0	0	0	0
3	0	0	0	0	0
4	0	0	0	0	0
5	0	0	0	0	0
6	1	1	0	2	0

Final error = 5.000 (2 solutions)

(b) Matrice immagine, Inconsistenze e Funzione criterio

Figura 5.12: *Two-mode* Blockmodeling generalizzato delle co-produzioni teatrali 2014-2015.

Complessivamente, nell'evoluzione della struttura, e delle posizioni e ruoli che

5.4 Analisi del campo attraverso l'Analisi Fattoriale Multipla e il Blockmodeling: il caso della produzione teatrale

la caratterizzano, sono visibili alcune variazioni nel tempo. Rispetto al modello complessivo di relazioni tra le posizioni - ossia, il blockmodel - si nota una stabilità del modello segmentato di fondo su cui, nel tempo, si innesca una tendenza verso la gerarchizzazione che porta ad assumere un modello *molti-core multi-periphery*. Nel susseguirsi delle stagioni, si assiste ad una lieve variazione nei modelli di connessioni. Infatti, per alcuni blocchi, la densità interna, e quindi il grado di similarità nei *pattern* relazionali che li contraddistinguono in termini di partecipazioni comuni, subisce una flessione a favore di una maggiore apertura verso la collaborazione tra posizioni (di compagnie) diverse. Quest'ultima viene segnalata dalla co-partecipazione di membri appartenenti a posizioni distinte nelle coproduzioni di uno stesso cluster. I modelli sembrano mostrare una struttura caratterizzata da una gerarchia duplice. Ad un livello di sistema complessivo, la prima gerarchia è definita dall'opposizione tra, da un lato, un'area del network caratterizzata da cluster di collaborazioni intense, associati a posizioni dominanti, che sembrano in grado di controllare e monopolizzare risorse di varia natura. Dall'alto, un'area del network caratterizzata da cluster di collaborazioni sparse, talvolta esclusive, che sono associati a posizioni periferiche (in un certo senso, dominate). Il secondo tipo di gerarchia ha una connotazione specifica e riguarda i rapporti di collaborazione, nelle stesse coproduzioni, tra le posizioni appartenenti a queste due aree del sistema. Tale modello implica il coinvolgimento sia di compagnie marginali, sia di organizzazioni rilevanti a livello nazionale (Teatro-Roma, StabileGen, StabileTos), il che pone determinati cluster in una posizione di mediazione tra aree interne del sistema teatrale regionale e tra quello regionale e nazionale con risvolti sulla possibilità di veicolare all'interno del sistema nuove risorse ed opportunità di varia natura. A livello specifico delle posizioni per le organizzazioni, emergono 3 posizioni chiave - Campania Festival e Stabile Napoli; ProgMuseo e Nuvoletti; Teatro Azione e Magazzini - la cui composizione rimane più o meno stabile nel tempo, ad eccezione di Campania Festival che nell'ultima stagione perde importanza assumendo un ruolo meno rilevante a livello strutturale. Ad essa se ne aggiunge una quarta - Bracco e Totò - che nell'ultima stagione scompare dalla struttura delle coproduzioni. Inoltre, nell'ultima stagione, è possibile notare l'emergere di un nuovo attore, ExAsilo.

5.4 Analisi del campo attraverso l'Analisi Fattoriale Multipla e il Blockmodeling: il caso della produzione teatrale

Le analisi di Blockmodeling sin qui proposte offrono una visione della struttura relazionale sottostante il network delle coproduzioni teatrali, con la definizione delle posizioni degli attori (e dei cluster di coproduzioni) e delle relazioni di equivalenza che le legano, fondate sui *pattern* di coproduzione che definiscono il loro ruolo nel network. Tuttavia, questi risultati, non mostrano le peculiarità legate alle caratteristiche degli attori e delle coproduzioni tra i quali intercorrono tali *pattern* relazionali, grazie alle quali è possibile ottenere una comprensione, ed una interpretazione, più profonda. Nel tentativo di perseguire tale obiettivo e unire la struttura relazionale e quella oggettiva per lo studio della struttura del campo si procede con l'integrazione dei risultati del Blockmodeling nel *frame* dell'analisi multidimensionale.

Tuttavia, prima di proseguire è necessario evidenziare che è stato condotto anche un tentativo parallelo di analisi dei network di affiliazione attraverso la realizzazione di una serie di Blockmodeling esplorativi secondo il criterio di equivalenza strutturale e regolare per tutte le stagioni teatrali. Tuttavia, data la complessità delle strutture dei network tale procedura si è rivelata incapace di identificare tutti i modelli di connessioni presenti al suo interno. Nel primo caso ha fornito solo indicazioni sull'esistenza di alcuni blocchi strutturalmente equivalenti, nel secondo ha restituito una molteplicità di partizioni ugualmente ben modellate ma non interpretabili.

5.4.4 Analisi Fattoriale Multipla e Blockmodeling: risultati

L'analisi della struttura di un campo ricostruita attraverso un network di affiliazione esprime una doppia dualità, una in termini di network, l'altra in termini di campo. La prima forma, in termini di network, è definita, nel presente caso, dalla relazione tra produttori (attori) e prodotti culturali (eventi), in quanto i produttori entrano in relazione e sono definiti a partire dalle coproduzioni a cui prendono parte, e, allo stesso tempo, le coproduzioni sono definite a partire dagli attori che comunemente le danno origine prendendo parte ad esse. In questa proposta, tale dualità è stata analizzata attraverso l'approccio posizionale e il Blockmodeling, individuando le classi di compagnie e coproduzioni equivalenti nella struttura relazionale sulla base dei loro *pattern* di relazioni.

5.4 Analisi del campo attraverso l'Analisi Fattoriale Multipla e il Blockmodeling: il caso della produzione teatrale

La seconda forma di dualità, in termini di campo, è data dallo spazio delle posizioni sociali (attori) e dallo spazio delle prese di posizione (eventi). Entrambi sono definiti dalla distribuzione di risorse che corrispondono agli attributi categoriali, rispettivamente, delle compagnie e delle coproduzioni (183). In questa proposta, tale dualità è analizzata attraverso l'approccio multidimensionale dell'AFM, e al suo interno dell'ACM, che consente di ricostruire le relazioni oggettive tra le posizioni di campo, ossia la collocazione dei soggetti nello spazio sociale definito dagli attributi, e contemporaneamente le prese di posizione, ossia le scelte manifeste dei soggetti occupanti una posizione, effettuate in virtù della loro collocazione, rispetto alle altre posizioni, nella struttura oggettiva. La dualità definita in tal modo implica la sovrapposizione di due strutture di natura diversa, una legata alle relazioni di network, e l'altra, legata alle relazioni oggettive.

In quest'ottica, se applicata soltanto agli attributi, l'ACM consente solo di valutare le relazioni tra le caratteristiche degli attori e degli eventi, cogliendo la struttura oggettiva, ignorando gli aspetti che riguardano le relazioni di network che intercorrono tra loro, ossia la struttura relazionale. La nostra proposta si innesta proprio su questo aspetto, in quanto, l'analisi diretta del network di affiliazione attraverso l'ACM e il Blockmodeling consente di tenere insieme entrambe le strutture, ancorando quella oggettiva (più astratta) a quella delle relazioni di network (più concrete). Essa permette di ricostruire la struttura delle relazioni sociali, e, attraverso l'introduzione degli attributi (degli attori e degli eventi), consente di ricostruire, sovrapponendola alla prima, anche la struttura delle relazioni oggettive. L'introduzione nell'analisi dei cluster equivalenti degli attori e degli eventi, ottenute attraverso il Blockmodeling, consente di mettere analizzare l'associazione diretta le posizioni di network con le posizioni (e prese di posizione) oggettive. Il risultato è l'individuazione di posizioni (per gli attori) e pre-di-posizioni (per gli eventi) definite in modo più completo sulla base di entrambe le componenti.

Il tutto viene inserito in un'ottica diacronica attraverso l'AFM, che consente di tenere insieme anche la dimensione dell'evoluzione del network.

In questi aspetti concettuali e metodologici l'approccio combinato di AFM e Blockmodeling, da noi proposto in questo lavoro, costituisce una novità nell'ambito degli approcci della SNA per l'analisi dei network di affiliazione.

5.4 Analisi del campo attraverso l'Analisi Fattoriale Multipla e il Blockmodeling: il caso della produzione teatrale

Date tali premesse, il passo successivo nella ricostruzione del campo delle coproduzioni teatrali è quello di inserire le classi equivalenti di compagnie e coproduzioni (posizioni di network), emerse dalla struttura relazionale ottenuta attraverso il Blockmodeling, in uno spazio sociale multidimensionale costruito a partire sia dalle relazioni di coproduzione sia dagli attributi (delle compagnie e delle coproduzioni) attraverso l'ACM e, più in generale, per coglierne l'evoluzione temporale, attraverso l'AFM. Questa strategia di analisi a tre vie, a nostro avviso innovativa, consente di applicare la logica "relazionale" all'analisi di un network di affiliazione che varia nel tempo, cogliendo sia la struttura sia la dinamica del campo.

In quest'ottica due compagnie che sono prossime nello spazio sociale definito dalla mappa fattoriale, sono in una relazione di equivalenza distribuzionale, ossia hanno *pattern* di partecipazione simili, il che significa anche che occupano la stessa posizione nel campo e condividono relazioni strutturali simili rispetto alle risorse economiche e culturali.

L'analisi del campo delle coproduzioni teatrali. L'analisi del campo delle coproduzioni attraverso l'AFM è stata condotta sulla *grand matrix* dall'affiancamento delle matrici di affiliazione delle quattro stagioni teatrali e composta da 100 compagnie e 157 coproduzioni. Nella matrice di affiliazione finale, gli attributi e le posizioni del Blockmodeling, entrambi degli attori e degli eventi, sono stati inseriti nell'analisi come variabili e punti supplementari (Paragrafo 3.4). Al fine di estrarre le informazioni racchiuse in tutti momenti temporali e individuare i fattori comuni si presentano risultati dell'analisi globale¹. La Tabella 5.2 presenta i risultati dell'analisi globale dell'AFM sulla base delle quattro stagioni di coproduzioni teatrali. I dati mostrano l'inerzia non rivalutata e rivalutata riprodotta dalle prime cinque dimensioni della soluzione fattoriale e forniscono una misura della qualità dell'approssimazione dei dati originali ottenuta attraverso la rappresentazione grafica basata su tali dimensioni. Come si può notare, le prime due dimensioni riproducono il 56% dell'inerzia totale. La mappa fattoriale costruita a partire da queste due dimensioni costituisce una buona approssimazione

¹Di seguito si riportano solo alcuni risultati relativi all'analisi globale. I risultati completi sono disponibili su richiesta.

5.4 Analisi del campo attraverso l'Analisi Fattoriale Multipla e il Blockmodeling: il caso della produzione teatrale

	λ	λ_{adj}	$In.adj.$	$In.adj.cum.$
Dim1	3,72	13,88	0,37	0,37
Dim2	2,69	7,26	0,19	0,56
Dim3	2,62	6,91	0,18	0,74
Dim4	1,89	3,58	0,09	0,83
Dim5	1,11	1,22	0,03	0,86
Dim6	0,95	0,90	0,02	0,89
Dim7	0,86	0,75	0,02	0,91
Dim8	0,81	0,65	0,02	0,93

Tabella 5.2: AFM. Inerzia rivalutata. Analisi globale.

	Dim1	Dim2	Dim3	Dim4	Dim5
11	0,90	0,80	0,84	0,79	0,20
12	0,94	0,69	0,67	0,64	0,28
13	0,90	0,60	0,65	0,45	0,42
14	0,98	0,60	0,45	0,01	0,21

Tabella 5.3: AFM. Coordinate delle stagioni sulle dimensioni dell'analisi globale.

della struttura dei dati originali, in quanto le distanze tra gli attori e gli aventi approssimano in modo ottimale le distanze presenti negli spazi originari (118). Pertanto, tale mappa sarà utilizzata per rappresentare e analizzare graficamente le similarità e le differenze tra le unità nella struttura del campo. Tuttavia, va evidenziato che anche la terza e la quarta dimensione riproducono un quota non irrilevante di inerzia, rispettivamente, il 18% e il 9% che unita a quella delle prime due conduce all'83% dell'inerzia totale, pertanto tali dimensioni anche se non saranno mostrate graficamente verranno tenute in considerazione nel corso di tutta l'analisi. Questo elemento è molto importante perchè consente di mostrare ed evidenziare un aspetto centrale emerso nel corso del quarto capitolo, legato alla capacità dell'analisi fattoriale di ricostruire sulle prime dimensioni la struttura forte del network, soprattutto le classi di unità strutturalmente equivalenti.

La scelta della rappresentazione bidimensionale è confermata anche dall'analisi delle coordinate delle singole stagioni sulle dimensioni dell'analisi globale riportare in Tabella 5.3. Tutte e quattro le stagioni assumono coordinate positive elevate sulle prime due dimensioni.

Rappresentazioni delle compagnie nello spazio globale delle coproduzioni. La Figura 5.13 rappresenta le compagnie nello lo spazio definito da tutti

5.4 Analisi del campo attraverso l'Analisi Fattoriale Multipla e il Blockmodeling: il caso della produzione teatrale

le coproduzioni, per tutte le stagioni teatrali. In questo spazio, ciascun punto rappresenta un attore, e le sue coordinate sono date dalla media ponderata delle coordinate nell'arco dei quattro anni. La localizzazione delle compagnie indicano la tendenza a collaborare in alcune coproduzioni specifiche nel corso del tempo. La prossimità tra i punti degli attori può essere interpretata in termini di similarità relazionale, in particolare in termini di similarità nei loro modelli di partecipazione (169). Pertanto, le compagnie coinvolte nelle stesse co-produzioni tendono ad essere collocate vicine sul piano fattoriale. In questo modo è possibile identificare facilmente gruppi di attori e attori più isolati.

Attraverso l'interpretazione degli assi e l'analisi delle coordinate dei punti (globali e parziali) (Figura 5.4), è possibile valutare i profili di partecipazione degli attori in base alla loro collocazione sul piano fattoriale. I punti collocati ad una maggiore distanza dall'origine che hanno le coordinate più elevate sull'asse presentano *pattern* di partecipazione più rari, mentre quelli collocati in prossimità dell'origine presentano quelli più comuni. Nel caso in esame, la maggiore distanza dall'origine degli assi di alcune compagnie, rappresentate dai punti presenti sul lato destro della mappa, segnala proprio la peculiarità dei loro *pattern* relazionali. Stabile Napoli, collocato nell'estremo superiore destro della mappa, occupa tale posizione in virtù della sua partecipazione costante nel corso delle stagioni a un certo numero di coproduzioni con Campania Festival, situata nella stessa area della mappa. Le due compagnie formano, quindi, un primo gruppo di soggetti con *pattern* simili (sono legate in gran parte alle stesse coproduzioni). Il secondo gruppo chiaramente individuabile è costituito da Nuvole, posizionato all'estremo destro inferiore della mappa. Nella stessa area si trova Prog Museo, che risulta fortemente legato a Nuvole nei quattro istanti temporali. Dall'analisi delle coordinate (Figura 5.4) dei primi due assi fattoriali è possibile osservare chiaramente che quelle più elevate sono assunte da questi due gruppi di attori.

In particolare, la Nuvole e Prog Museo sono le compagnie che assumono coordinate positive elevate solo sul primo asse, mentre Stabile e Campania Festival le presentano sul primo e sul secondo asse. Tale collocazione suggerisce che l'interpretazione del primo asse è fortemente legata ai modelli di partecipazione di questi due gruppi e in particolare alla loro collaborazione su specifiche coproduzioni nel corso dei quattro anni. Sul secondo fattore, invece, le coordinate

5.4 Analisi del campo attraverso l'Analisi Fattoriale Multipla e il Blockmodeling: il caso della produzione teatrale

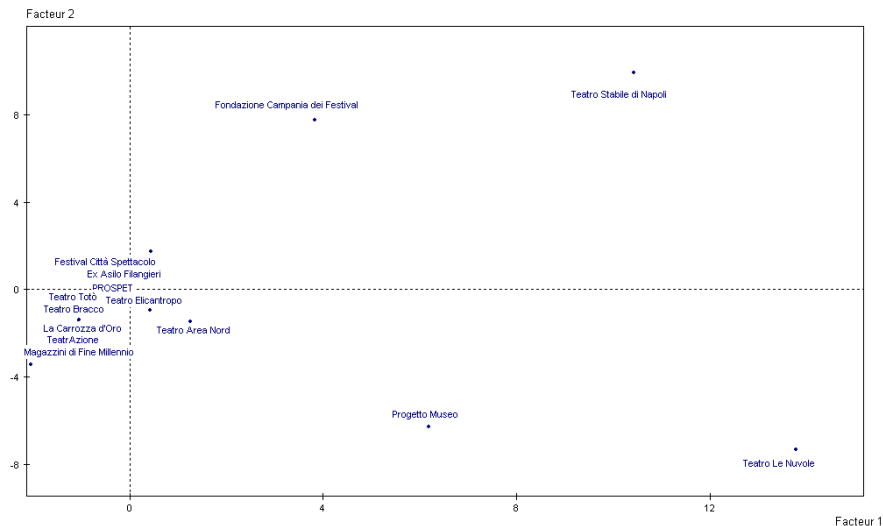


Figura 5.13: Analisi globale: rappresentazione delle compagnie nello spazio delle coproduzioni attraverso le coordinate principali. Ogni punto rappresenta un attore e le sue coordinate sono la media ponderata delle coordinate assunte nelle quattro stagioni. Gli attributi delle compagnie sono aggiunti come punti supplementari. Proporzione di inerzia: primo fattore = 37%, secondo fattore 19%.

positive più alte nel corso dei quattro anni sono date da quelle del gruppo di compagnie Stabile e Campania Festival, a definire questo asse come caratterizzato soprattutto dalla collaborazione esclusiva tra questi due attori. Oltre questi due gruppi particolarmente rilevanti, è possibile individuarne altri due, collocati nel quadrante inferiore sinistro, che presentano *pattern* di collaborazioni opposti ai gruppi situati sul lato destro, tale opposizione è segnalata dalle coordinate negative più elevate assunte dalle singole compagnie sui due primi assi fattoriali nel corso dei quattro anni. Si tratta soprattutto della collaborazione tra Magazzini, Teatro Azione e Carrozza e di quella tra Bracco e Totò. In prossimità dell'origine degli assi, infine, è possibile individuare un quinto insieme di compagnie caratterizzato da *pattern* di collaborazione diffusa, non legata a una o più compagnie che partecipano frequentemente alle stesse coproduzioni. In altri termini, le imprese raggruppate in questa area hanno un tasso medio di partecipazione (definito a partire dal numero di coproduzioni alle quali partecipano) molto basso, e la loro partecipazione al sistema di coproduzioni della regione è sparsa e poco rilevante.

5.4 Analisi del campo attraverso l'Analisi Fattoriale Multipla e il Blockmodeling: il caso della produzione teatrale

	Dim.1	Dim.2	Dim.3	Dim.4	Dim.5
	<i>15,90</i>	<i>11,51</i>	<i>11,23</i>	<i>8,08</i>	<i>4,73</i>
Nuvole	13,79	-7,33	-0,34	0,25	0,16
ProgMuseo	6,19	-6,28	-0,84	-0,21	-0,77
StabileNap	10,43	9,92	2,52	0,84	0,45
CampaniaFestival	3,82	7,75	1,83	0,23	-0,47
Magazzini	-1,65	-2,49	12,74	0,82	0,12
TeatrAzione	-1,19	-1,70	7,87	0,47	0,15
Carrozza	-0,85	-0,95	4,06	0,09	-0,06
Totò	-0,92	-0,35	-1,29	10,22	-0,17
Bracco	-0,81	-0,29	-1,10	8,67	-1,21
PROSPET	-0,58	-0,22	-0,65	0,93	8,03
Elicantropo	-0,19	-0,43	-0,51	0,00	6,19
<i>Coordinate parziali</i>					
Nuvole.11	12,21	-7,09	-0,51	0,09	1,52
Nuvole.12	12,95	-5,43	-0,07	0,25	-0,10
Nuvole.13	15,10	-9,32	-0,62	0,11	-0,56
Nuvole.14	14,90	-7,47	-0,18	0,56	-0,21
ProgMuseo.11	6,71	-6,78	-0,95	-0,33	-0,75
ProgMuseo.12	4,98	-5,20	-0,76	-0,29	-0,64
ProgMuseo.13	6,76	-6,69	-0,86	-0,23	-0,93
ProgMuseo.14	6,33	-6,45	-0,78	0,03	-0,75
StabileNap.11	10,31	11,77	2,89	0,78	0,04
StabileNap.12	12,33	11,23	2,99	1,03	0,69
StabileNap.13	7,60	6,74	1,65	0,49	0,48
StabileNap.14	11,49	9,94	2,56	1,04	0,59
CampaniaFestival.11	6,21	11,83	2,47	0,34	-0,60
CampaniaFestival.12	4,80	9,58	1,96	0,26	-0,50
CampaniaFestival.13	2,31	5,41	2,10	0,05	-0,64
CampaniaFestival.14	1,98	4,18	0,79	0,28	-0,14
Magazzini.11	-2,05	-3,42	16,33	1,01	0,34
Magazzini.12	-1,77	-2,77	13,17	0,74	0,05
Magazzini.13	-1,47	-1,94	13,05	0,83	-0,06
Magazzini.14	-1,32	-1,81	8,41	0,70	0,14
TeatrAzione.11	-1,38	-2,16	10,00	0,54	0,29
TeatrAzione.12	-1,16	-1,69	7,54	0,37	0,15
TeatrAzione.13	-1,13	-1,53	7,55	0,41	0,03
TeatrAzione.14	-1,11	-1,41	6,40	0,55	0,12
Carrozza.11	-1,04	-1,37	5,97	0,12	0,02
Carrozza.12	-0,90	-1,09	4,49	0,05	-0,06
Carrozza.13	-0,82	-0,83	3,98	0,06	-0,18
Carrozza.14	-0,65	-0,49	1,80	0,15	-0,05
Toto.11	-1,14	-0,50	-1,86	16,13	-2,09
Toto.12	-1,07	-0,49	-1,63	13,68	0,12
Toto.13	-1,03	-0,30	-1,47	11,09	1,35
Toto.14	-0,44	-0,09	-0,22	0,00	-0,06
Bracco.11	-1,14	-0,50	-1,86	16,13	-2,09
Bracco.12	-0,97	-0,43	-1,45	12,20	-1,66
Bracco.13	-0,68	-0,14	-0,87	6,37	-1,04
Bracco.14	-0,44	-0,09	-0,22	0,00	-0,06

Tabella 5.4: AFM. Coordinate e coordinate parziali delle compagnie con contributi elevati.

Rappresentazione delle coproduzioni nello spazio globale delle compagnie. La Figura 5.14 rappresenta tutte le coproduzioni nello spazio definito dalle compagnie, per tutte le stagioni teatrali. In questo spazio, ogni evento è rappresentato da una coppia di punti collegati da un segmento. In esso possono essere analizzate visivamente le similarità, le associazioni, e cambiamenti tra i modelli di partecipazione ai diversi eventi e nei vari istanti temporali. Nello spazio

5.4 Analisi del campo attraverso l'Analisi Fattoriale Multipla e il Blockmodeling: il caso della produzione teatrale

degli attori è possibile individuare due elementi importanti come guida dell'analisi: *i*) l'angolo tra due vettori di eventi, che è proporzionale alla correlazione tra i *pattern* di partecipazione all'evento; e *ii*) la lunghezza dei vettori degli eventi, che è legata alla variabilità nel *pattern* di partecipazione (71, 169). Osservando gli angoli tra i profili di partecipazione all'evento per le diverse stagioni, è possibile valutare graficamente la presenza di alcuni gruppi di eventi caratterizzati da un'elevata similarità relazionale, in quanto l'angolo formato dai vettori risulta di dimensioni ridotte. Nel caso in esame, è possibile esaminare i *pattern* relazionali delle coproduzioni nello spazio delle compagnie. Come è visibile dalla mappa fattoriale, in alcuni casi, i segmenti che rappresentano le coproduzioni sono praticamente sovrapposti, il che indica che ad un certo numero di coproduzioni, aventi luogo in più stagioni teatrali, hanno partecipato le stesse compagnie. In questa mappa si può notare come il raggruppamento delle coproduzioni rifletta quello delle compagnie. In essa possono essere individuati quattro gruppi principali di coproduzioni con *pattern* relazionali simili tra loro, ma in netta opposizione tra gruppo e gruppo.

Un primo gruppo, collocato nel quadrante superiore destro, è composto dalle coproduzioni realizzate dal gruppo Stabile Napoli e Campania Festival, in collaborazione con poche altre compagnie. Il secondo gruppo è situato nel quadrante destro in basso, e rappresenta gli spettacoli co-prodotti da Nuvole e Prog Museo. Un terzo e un quarto gruppo corrispondono, invece, alle coproduzioni del terzo e quarto gruppo di compagnie evidenziate dall'analisi globale nello spazio degli eventi, ossia Magazzini, Teatro Azione e Carrozza; Bracco e Totò. È necessario notare che i *pattern* di partecipazione di questi gruppi si possono apprezzare meglio sulla terza e sulla quarta dimensione della soluzione fattoriale dove, essi, presentano i contributi positivi più elevati e dove le coproduzioni ad assi associati risultano maggiormente caratterizzanti (Figura 5.5).

Inoltre, occorre evidenziare che le coproduzioni dei primi tre gruppi, e in particolare quelle situate sul lato destro della mappa, sono molto "polarizzate", ossia caratterizzate da *pattern* di partecipazione piuttosto rari, la parte di segmento che dall'origine degli assi procede verso l'estremo destro della mappa è molto più lunga di quella situata nell'altro estremo. Si tratta di spettacoli che hanno coinvolto la partecipare in molte occasioni sempre delle stesse compagnie. Inoltre, si

5.4 Analisi del campo attraverso l'Analisi Fattoriale Multipla e il Blockmodeling: il caso della produzione teatrale

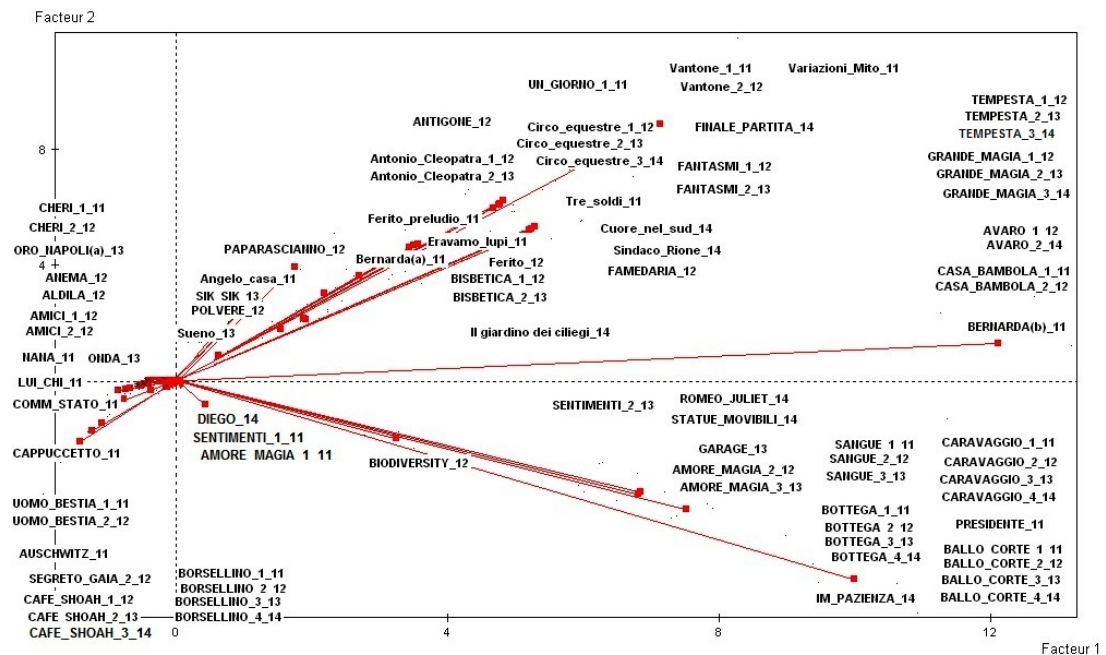


Figura 5.14: rappresentazione delle coproduzioni nello spazio delle compagnie per tutte le stagioni attraverso le coordinate principali. Ogni coproduzione è rappresentata da una coppia di punti connessi da un segmento. Proporzione di inerzia: primo fattore = 37%, secondo fattore 19%.

può notare come gran parte delle coproduzioni del primo gruppo siano negativamente correlate con quelle del terzo e del quarto gruppo, ossia chi ha partecipato alle coproduzioni di un gruppo non ha partecipato a quelle degli altri due. Nei termini dell'analisi del campo di Bourdieu, questo significa che la posizione di un gruppo è in una relazione di opposizione con quella di un gruppo situato nell'area opposta dello spazio sociale, alla quale, potenzialmente, può essere associata anche una collocazione opposta nel *range* delle opportunità date dalla struttura del campo. Infine, occorre osservare come vi siano delle "intersezioni". Infatti, come segnalato nel caso della rappresentazione delle compagnie, le coordinate positive più elevate sul primo fattore sono quelle associate ai gruppi Stabile Napoli-Campania Festival e Nuvole-Prog Museo. Pertanto, una prima intersezione è data da alcune coproduzioni, rappresentate dallo stesso segmento (Bernarda.b; Casa.bambola, ecc.) che coinvolgono una compagnia del primo e una del secondo

5.4 Analisi del campo attraverso l'Analisi Fattoriale Multipla e il Blockmodeling: il caso della produzione teatrale

Dim1		Dim2	
Coprod. ($R^2 > 0.5$)	R^2	Coprod. ($R^2 > 0.4$)	R^2
BERNARDAb.11	0,81	Vantone.1.11	0,59
CASA.BAMBOLA.1.11	0,81	UN.GIORNO.1.11	0,59
AVARO.1.12	0,81	Variazioni.Mito.11	0,59
GRANDE.MAGIA.1.12	0,81	Vantone.2.12	0,59
TEMPESTA.1.12	0,81	UN.GIORNO.2.12	0,59
CASA.BAMBOLA.2.12	0,81	FINALE.PARTITA.14	0,59
GRANDE.MAGIA.2.13	0,81	Circo.equestre.1.12	0,44
TEMPESTA.2.13	0,81	Circo.equestre.2.13	0,44
TEMPESTA.3.14	0,81	Circo.equestre.3.14	0,44
AVARO.2.14	0,81	Tre.soldi.11	0,43
GRANDE.MAGIA.3.14	0,81	Antonio.Cleopatra.1.12	0,43
BALLO.CORTE.1.11	0,55	Antonio.Cleopatra.2.13	0,43
Dim3		Dim4	
Coprod. ($R^2 > 0.8$)	R^2	Coprod. ($R^2 > 0.8$)	R^2
SEGRETO.GAIA.1.11	0,83	AMICI.1.11	0,97
UOMO.BESTIA.1.11	0,83	CHERI.1.11	0,97
BORSELLINO.1.11	0,83	COMM.STATO.11	0,97
AUSCHWITZ.11	0,83	LUI.CHI.11	0,97
CAFE.SHOAH.1.12	0,83	NANA.11	0,97
SEGRETO.GAIA.2.12	0,83	ALDILA.12	0,97
UOMO.BESTIA.2.12	0,83	ANEMA.12	0,97
BORSELLINO.2.12	0,83	AMICI.2.12	0,97
CAFE.SHOAH.2.13	0,83	CHERI.2.12	0,97
UOMO.BESTIA.3.13	0,83	GIACCA.PAGLIETTA.12	0,97
BORSELLINO.3.13	0,83	ORO.NAPOLI.a.13	0,97
CAFE.SHOAH.3.14	0,83	ONDA.13	0,97

Tabella 5.5: Risultati dell'AFM. Coproduzioni significativamente associate alle prime quattro dimensioni.

gruppo (Teatro Stabile di Napoli e Le Nuvole), ossia i membri dei due gruppi che maggiormente caratterizzano il primo asse fattoriale. Questo *pattern* di partecipazione costituisce uno di quelli più inusuali. Infatti, considerando la lunghezza del segmento, essa è maggiormente accentuata verso il lato destro della mappa fattoriale. Una seconda intersezione è visibile in prossimità dell'origine degli assi e riguarda gli spettacoli che Campania Festival ha realizzato con una serie di compagnie che, nella rappresentazione delle compagnie, erano visibili in prossimità dell'origine degli assi, nel quadrante superiore destro.

Da una prima osservazione delle mappe fattoriali, si può, dunque, rilevare che: (i) Il sistema delle coproduzioni teatrali in Campania, nel periodo oggetto di studio, risulta segmentato in base ai rapporti di collaborazione tra le compagnie e

(ii) La distinzione tra chi collabora a quale coproduzione sembra evidenziare una gerarchia basata sulla concentrazione di attività produttiva nelle mani di alcune compagnie.

Rappresentazione congiunta di compagnie e coproduzioni. Al fine di apprezzare le relazioni tra gli attori e gli eventi preservandone la naturale asimmetria si rappresentano entrambi in una mappa bidimensionale congiunta attraverso il biplot (108), nello specifico in una mappa asimmetrica in cui gli attori sono rappresentati in coordinate principali e gli eventi in coordinate standard. In questo caso l'interpretazione delle posizioni, intese come localizzazione, degli attori rispetto agli eventi è ottenuta utilizzando il vettore direzione definito da ciascun evento, che rappresenta gli assi del biplot. Proiettando i punti che rappresentano gli attori su ogni asse del biplot, possiamo ottenere un valore approssimativo del profilo di partecipazione all'evento che consente di ottenere delle intuizioni sul motivo per cui gli attori sono collocati vicini o lontani nello spazio degli eventi (169). Osservando la Figura 5.15 è possibile evidenziare la corrispondenza tra i gruppi evidenziati nelle singole rappresentazioni delle compagnie e delle coproduzioni. In particolare possibile identificare un primo dato da Stabile Napoli e Campania Festival (nel quadrante in alto a destra) che hanno le più elevate coordinate sugli assi biplot determinati dalla partecipazione congiunta alle coproduzioni Vantone; Variazione.Mito; Un.Giorno,ecc Questi attori sono caratterizzati da una elevata partecipazione a queste coproduzioni lungo tutto l'arco temporale considerato. Un secondo gruppo è dato da Nuvoles e Prog Museo che hanno le più elevate coordinate sugli assi biplot determinati dalla partecipazione congiunta alle coproduzioni Ballo.cortese; Presidente; Amore.magia, ecc.

Un terzo gruppo è dato da Teatro Azione e Magazzini che hanno le più elevate coordinate sugli assi biplot determinati dalla partecipazione congiunta alle coproduzioni Mulini.vento; Borsellino; Cappuccetto, ecc. Un quarto gruppo è dato da Bracco e Totò che hanno le più elevate coordinate sugli assi biplot determinati dalla partecipazione congiunta alle coproduzioni Ti.presento; Amici. Infine è possibile evidenziare i due gruppi trasversali rispetto ai gruppi definiti sino ad ora. In particolare il primo è dato dai teatri Stabile Napoli e Nuvoles che hanno le più elevate coordinate sugli assi biplot determinati dalla partecipazione

5.4 Analisi del campo attraverso l'Analisi Fattoriale Multipla e il Blockmodeling: il caso della produzione teatrale

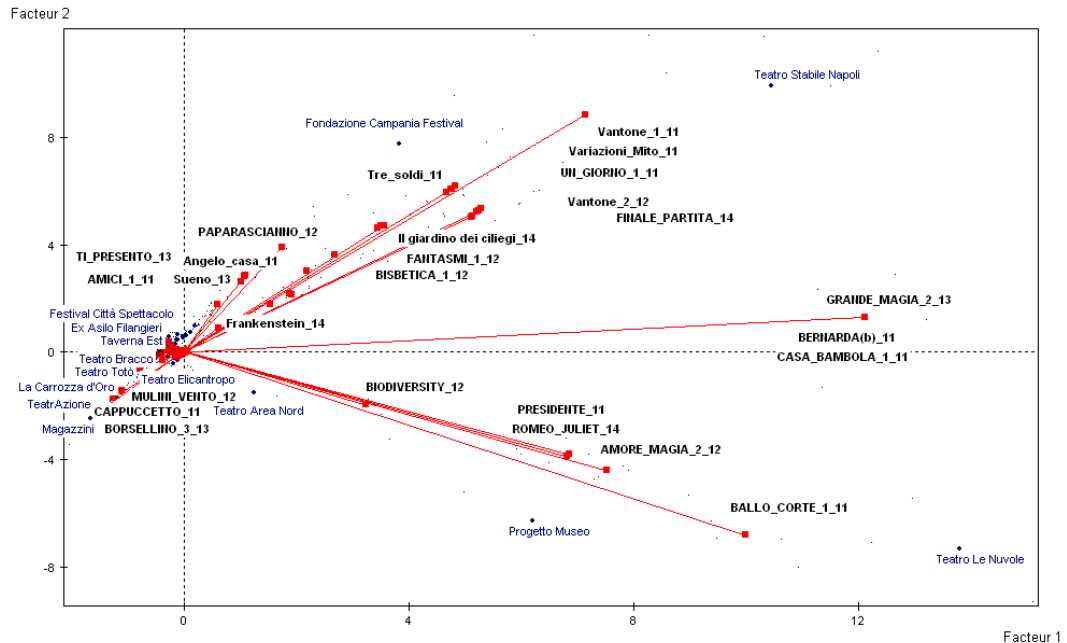


Figura 5.15: [Analisi globale: rappresentazione congiunta delle compagnie e delle coproduzioni. Biplot asimmetrico delle compagnie (punti) e delle coproduzioni (quadrati) con alcune compagnie e coproduzioni evidenziate. Le compagnie sono rappresentate in coordinate principali e le coproduzioni in coordinate standard. Proporzione di inerzia: primo fattore = 37%, secondo fattore 19%.

congiunta alle coproduzioni Bernarda.b; Casa.bambola, ecc. Mentre il secondo dato da Campania Festival e alcune compagnie collocate in prossimità dell'origine degli assi che hanno le più elevate coordinate sugli assi biplot determinati dalla partecipazione congiunta alle coproduzioni Parascianno, Angelo.casa; Sueno.

Variazioni nel tempo e traiettorie delle compagnie Al fine di analizzare le variazioni nel tempo della struttura relazionale globale, Tabella 5.6 *RV* riporta i valori dei coefficienti di *RV* che informa sul grado di similarità nella struttura del network, tra i diversi momenti temporali (169). I valori mostrano che la struttura relazione nel corso degli anni non subisce considerevoli cambiamenti. Infatti, considerato che il coefficiente assume valori *RV* tra 0 e 1, dove valori

5.4 Analisi del campo attraverso l'Analisi Fattoriale Multipla e il Blockmodeling: il caso della produzione teatrale

	11	12	13	14
11	1	-	-	-
12	0,88	1	-	-
13	0,77	0,84	1	-
14	0,70	0,75	0,78	1

Tabella 5.6: AFM. Coefficiente di similarità RV per le quattro stagioni stagioni di coproduzioni.

prossimi a 1 corrispondono a tabelle molto simili, nel caso in esame si osservano valori tutti al di sopra di 0.7. Dove, peraltro, tale valore corrisponde all'ultima stagione che risulta quella che si differenzia maggiormente. Questi risultati sono in linea con quelli osservati nell'analisi di Blockmodeling, in cui emergeva un cambiamento graduale più accentuato nella quarta stagione. Questo leggero e graduale cambiamento nei modelli di coproduzione nel corso degli anni si nota dal fatto che i valori relativi ad anni successivi sono più elevati (ad es. 2011-2012: 0,88; 2012-2013: 0,84, ecc rispetto a 2011-2014: 0,70).

Nello spazio globale delle coproduzioni è possibile osservare la variazione nelle traiettorie delle compagnie, che forniscono una visualizzazione combinata dei risultati dell'analisi parziale spazio globale delle coproduzioni sottolineando la dimensione temporale in termini di evoluzione. I segmenti che collegano i punti relativi allo stesso coproduzione nel corso dell'analisi parziale raffigurano la traiettoria della compagnia nelle quattro stagioni teatrali e forniscono una visualizzazione grafica dell'evoluzione individuale modello di partecipazione. La Figura 5.16 traccia le traiettorie temporali delle compagnie Campania Festival, Stabile Napoli, Nuvole. La prima compagnia presenta una traiettoria che partendo dall'area caratterizzata da una maggiore intensità di partecipazione si sposta gradualmente in direzione del centro, caratterizzata da modelli di partecipazione meno intensi. Questo cambiamento molto forte è stato riscontrato anche nell'analisi di Blockmodeling, dove, nell'ultima stagione, tale attore, sulla base del suo modello di partecipazione, risultava collocato in un blocco di collaborazioni sparse. Nella traiettoria di Stabile Napoli è possibile ravvisare un certo cambiamento, in direzione del centro, solo per la stagione 2013, infatti, in quella successiva il suo modello di partecipazione si ricolloca nell'area caratterizzata da collaborazione più intense. Un confronto tra le traiettorie delle due compagnie

5.4 Analisi del campo attraverso l'Analisi Fattoriale Multipla e il Blockmodeling: il caso della produzione teatrale

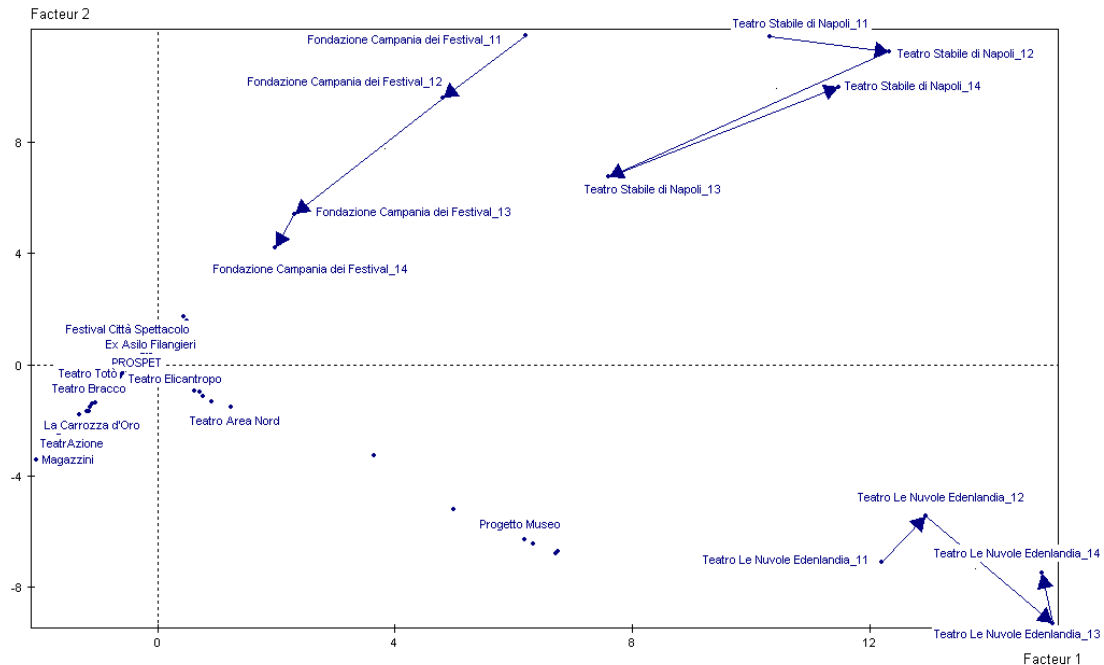


Figura 5.16: Analisi globale: rappresentazione delle compagnie nello spazio delle coproduzioni. Ogni punto rappresenta un attore in un dato momento. Sono evidenziate le traiettorie relative ad alcune compagnie. Proporzione di inerzia: primo fattore = 37%, secondo fattore 19%.

rende molto evidente la variazione tra i modelli di partecipazione dei due attori nell'ultima stagione, in cui si evidenzia una perdita di rilevanza di Campania Festival rispetto a Stabile Napoli che, invece, mantiene il suo ruolo dominante all'interno del sistema delle coproduzioni. Per quanto riguarda Nuvole, tale compagnia presenta una traiettoria con un direzione che nel corso del tempo di sposta in modo accentuato verso il quadrante in basso a destra della mappa fattoriale, a segnalare un modello di partecipazione molto intenso e che gli consente di mantenere un ruolo importante all'interno del sistema.

Rappresentazioni delle posizioni compagnie e degli attributi nello spazio delle coproduzioni globale. Infine, uno degli aspetti centrali dell'approccio proposto è costituito dalla possibilità di integrare nell'analisi, come variabili

5.4 Analisi del campo attraverso l'Analisi Fattoriale Multipla e il Blockmodeling: il caso della produzione teatrale

supplementari, le posizioni delle compagnie e gli attributi delle compagnie e delle coproduzioni in un'unica rappresentazione nello spazio delle coproduzioni. La rappresentazione ottenuta, e riportata in Figura 5.17, preserva tutte le informazioni contenute nella matrice di affiliazione originaria e consente di osservare meglio le caratteristiche della struttura di network già evidenziate sin ora, permettendo di esplorare l'associazione tra i modelli relazionali, le posizioni di network delle compagnie, e la struttura relazionale e degli attributi in cui esse sono rappresentate. Le etichette delle posizioni (caratterizzate dal colore nero) sono dotate di due valori numerici che corrispondono, rispettivamente, al cluster degli attori e al cluster degli eventi a cui è associato nel blockmodel, in questo modo è possibile non solo identificare la posizione degli attori ma anche precisamente il cluster di eventi a cui quella posizione è associata e, di conseguenza il modello di connessioni che lo ha caratterizzato nella definizione dei blocchi empirici. Si ricorda che ogni blocco empirico è conforme ad un certo tipo di equivalenza che riflette uno specifico modello di connessioni.

L'adozione di criteri di equivalenza aggiuntivi rispetto a quella strutturale ha consentito il raggruppamento di compagnie che, pur non avendo modelli relazionali identici, sono uniti dall'avere modelli di partecipazione in linea con specifici tipi di equivalenza (ad esempio l'essere stati coinvolti in almeno una coproduzione induce una forma di equivalenza regolare che rende tali attori regolarmente equivalenti, rispetto alle coproduzioni a cui sono associate). Questi elementi consentono di considerare anche i diversi ruoli che le compagnie esercitano nel sistema delle coproduzioni. Nell'analisi visuale della mappa fattoriale in esame è possibile valutare la collocazione delle posizioni e caratterizzarle osservando l'associazione delle stesse con gli attributi degli attori e degli eventi. La struttura del network emersa dal Blockmodeling, caratterizzata da segmentazione e gerarchizzazione riconducibili ad un modello misto tra sottogruppo coeso e *multi-core multi-periphery* è, in parte, visibile anche sul piano fattoriale. In particolare, emerge l'aspetto della segmentazione, dove è ben visibile la distinzione tra posizioni, formate dalle compagnie con *pattern* di relazione più simili emerse già nell'analisi delle singole rappresentazioni dell'AFM. In questo, modo possono essere ben distinti i quattro gruppi principali, che corrispondono ad altrettante posizioni, ed insieme ad

5.4 Analisi del campo attraverso l'Analisi Fattoriale Multipla e il Blockmodeling: il caso della produzione teatrale

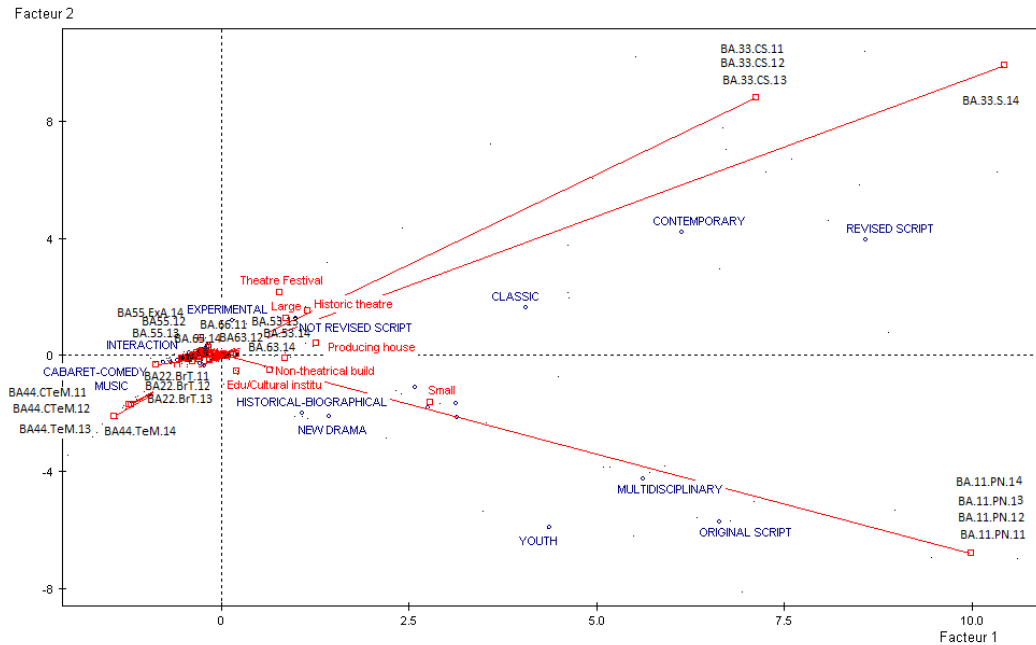


Figura 5.17: Analisi globale: rappresentazione delle posizioni delle compagnie e degli attributi delle compagnie e delle coproduzioni nello spazio delle coproduzioni. I cluster e gli attributi sono aggiunti come punti supplementari e sono rappresentate come medie dei punti degli attori in coordinate principali. Etichette: nero=Posizioni di network compagnie; rosso= attributi compagnie; blu= attributi coproduzioni. Proporzione di inerzia: primo fattore = 37%, secondo fattore 19%.

essi, sono visibili alcuni gruppi, in prossimità degli assi, caratterizzati da modelli di connessioni particolari che, generalmente, non sarebbe possibile visualizzare (B55.12; B55.13; BA55.ExA.14). Questo non sorprende poichè, come si è visto nel corso del terzo capitolo (Paragrafo 3.5), l'equivalenza distribuzionale, che caratterizza la metrica del χ^2 che definisce le distanze di questo spazio, sostanzialmente tende a riprodurre una configurazione relazionale molto simile a quella indotta da un Blockmodeling basato sull'equivalenza strutturale. Pertanto, le compagnie con modelli di partecipazione uguali, ma anche simili, tendono ad essere quelle più facilmente individuate nell'ACM che, nel caso di *pattern* stabili nel tempo come in questo caso, emergono anche nell'AFM. L'aspetto legato alla gerarchia,

5.4 Analisi del campo attraverso l'Analisi Fattoriale Multipla e il Blockmodeling: il caso della produzione teatrale

in particolare al *core-periphery*, è meno evidente. Tuttavia, con l'aiuto delle etichette è possibile individuare facilmente le posizioni che sono associate alle stesse coproduzioni, ad esempio è possibile notare la posizione *core* stabile nelle quattro stagioni, costituita da Campania Festival e Stabile Napoli (BA33.CS.11,12,13), in alto a destra della mappa, e le sue *periphery*, che variano nel tempo e nella composizione, costituite da alcune delle compagnie collocate in prossimità dell'origine degli assi (BA53.13; BA53.14; BA63.12). Inoltre, è possibile osservare quella esclusiva per la quarta stagione con Stabile Napoli (BA33.S.14), come *core* e alcune compagnie della collaborazione sparsa come *periphery* data dalla posizione BA63.4. Infine si può notare anche l'altro piccolo *core* emerso nell'ultima stagione, costituito da ExAsilo (BA55.ExA.14) con la sua *periphery* (BA65.14). Tale modello è molto interessante per le implicazioni a livello strutturale, infatti è associato alla possibilità che compagnie che tra loro collaborano poco o per nulla, abbiano modo di entrare in relazione attraverso le compagnie presenti nel gruppo coeso, che funge dunque da mediatore. Il caso di Teatro Stabile Napoli e Campania Festival è appunto questo: si tratta di soggetti con una forte posizione in termini istituzionali, e con una dotazione consistente di finanziamenti. Per sua natura, inoltre, un festival è fatto per operare in sinergia con il territorio e agisce, in questo caso, anche in modo tale da aprire spazi di azione a soggetti in qualche modo marginali, come quelli situati nelle posizioni caratterizzate da una intensità di collaborazione inferiore. Le posizioni così rappresentate possono essere analizzate anche osservando la collocazione relativa nello spazio fattoriale, in questo caso essa ricalca in modo sintetico quella delle compagnie al loro interno, collocate nello spazio in relazione ai modelli di partecipazione alle coproduzioni. Questa, rappresentazione, tuttavia, in un'ottica di Blockmodeling consente di visualizzare le posizioni in uno spazio metrico in cui le distanze assumono pieno significato. Inoltre, considerando gli attributi delle compagnie e delle coproduzioni, tali posizioni vengono inserite in un contesto significativo in cui possono essere caratterizzate e interpretate, unendo la componente relazionale a quella categoriale. Nel dettaglio, nella Figura 5.7 sono riportati i risultati relativi ai cluster di attori che, nel corso delle stagioni, sono associate in modo significativo alle prime quattro dimensioni della soluzione fattoriale e i relativi attributi. Inoltre,

5.4 Analisi del campo attraverso l'Analisi Fattoriale Multipla e il Blockmodeling: il caso della produzione teatrale

Dim1		Dim2		Dim3		Dim4	
Cluster Com e att.	R^2	Cluster Com e att.	R^2	Cluster Com e att.	R^2	Cluster Com e att.	R^2
BA33.CS.11	0,59	BA11.PN.11	0,55	BA44.TeM.14	0,83	BA22.BrT.11	0,97
BA33.CS.12	0,59	BA11.PN.12	0,55	BA44.CTeM.11	0,80	BA22.BrT.12	0,97
BA33.CS.13	0,59	BA11.PN.13	0,55	BA44.CTeM.12	0,80	BA22.BrT.13	0,97
BA33.S.14	0,37	BA11.PN.14	0,55	BA44.TeM.13	0,24	SC.Large	0,22
BA11.PN.11	0,35	BA33.S.14	0,30	SC.Medium	0,08	ST.Modern.th	0,13
BA11.PN.12	0,35	BA33.CS.11	0,28			Nod.typ.Foc	0,09
BA11.PN.13	0,35	BA33.CS.12	0,28			T.Build.bas.th.comp	0,08
BA11.PN.14	0,35	BA33.CS.13	0,28				
ST.Hist.theat	0,07	SC.Small	0,11				
T.Theat.Fest	0,07	T.Prod.h	0,09				
SC.Large	0,06	Nod.typ.Foc	0,05				
SC.Small	0,05						

Tabella 5.7: Risultati dell'AFM con cluster delle compagnie e attributi. Cluster delle compagnie e attributi maggiormente associate alle prime cinque dimensioni.

nella Figura 5.8 sono riportati i risultati relativi alle coordinate parziali dei cluster delle coproduzioni e le coordinate globali degli attributi che maggiormente caratterizzano le prime cinque dimensioni della soluzione fattoriale.

La possibilità di considerare gli attributi rappresenta anche un aspetto specifico dell'analisi del campo, perchè attraverso le caratteristiche delle coproduzioni quali, nel caso in esame, il tipo di testo/rappresentazione e i generi teatrali degli spettacoli realizzati, è possibile osservare anche le prese di posizione di questi insiemi di compagnie. Di conseguenza, nel caso in esame, dalla mappa fattoriale è possibile osservare come i modelli di partecipazione delle due posizioni collocate a destra, almeno per i primo tre anni, si differenzino notevolmente da quelli collocati nel lato sinistro e in prossimità degli assi. Evidenziando le caratteristiche che riflettono tale differenziazione. Infatti, è possibile osservare come la prima posizione occupata da Campania Festival e Stabile Napoli (BA.33.CS) sia composta da membri appartenenti organizzazioni importanti del sistema teatrale (una Fondazione che gestisce un festival teatrale e un teatro Stabile con riconoscimento istituzionale), con l'accesso a risorse economiche di varia natura che gli consente la produzione diretta degli spettacoli (in quanto riconosciute possono disporre di finanziamenti statali (*producing house*)), inoltre, nel caso di Stabile Napoli, è dotato di una struttura di grandi dimensioni che rispetta i canoni spaziali che storicamente caratterizzano l'edificio del teatro (182). Dal punto di vista delle caratteristiche della produzione teatrale, questa posizione è associata a coproduzioni che coinvolgono testi letterari contemporanei e classici, perlopiù rivisitati (*revised script*). Tali caratteristiche portano ad interpretare tale posizione come

5.5 Analisi Fattoriale Multipla e Blockmodeling: considerazioni finali

(BA.22) caratterizzata anch'essa dall'essere composta da imprese di produzione, dotate di strutture relativamente grandi che realizza prevalentemente spettacoli di teatro comico dialettale o cabaret, con una pronunciata vocazione all'intrattenimento. Un quinto insieme di posizioni è ancorato all'origine degli assi; si tratta di posizioni misti di compagnie dedite alla sperimentazione, al dramma storico biografico o alla nuova drammaturgia. È possibile notare, in un certo senso, che le compagnie più marginali in termini relazionali lo sono anche sul piano della produzione teatrale. Questo, infatti può essere considerato come il versante dominato della struttura di coproduzione che per accedere a maggiori risorse deve cercare di costruire legami di coproduzione con le posizioni del polo dominante.

In conclusione, i processi di segmentazione e gerarchia presenti nel network sono legati a una concentrazione di attività nelle mani di alcune compagnie che, da un lato, stabiliscono intensi rapporti di collaborazione tra loro, duraturi nel tempo, e dall'altro coinvolgono occasionalmente altri soggetti esterni ai principali gruppi di riferimento. I generi *mainstream* sono quelli del teatro classico e contemporaneo, praticato dal blocco "dominante", che manifesta una struttura di tipo *core-periphery*, se considerato in relazione ai blocchi misti i quali, per la guadagnare visibilità e contributi finanziari intrattengono legami di collaborazione con il *core* di riferimento. I generi di intrattenimento sono associati invece a una segmentazione, o vero e proprio isolamento, ma anche a una netta opposizione sia relazionale che in termini di genere rispetto al resto del network.

5.5 Analisi Fattoriale Multipla e Blockmodeling: considerazioni finali

In conclusione, è possibile sostenere che l'approccio combinato dell'AFM e del Blockmodeling risulta particolarmente fruttuosa per entrambi. Il Blockmodeling fornisce una partizione simultanea, ma distinta, delle righe e delle colonne della matrice di affiliazione, consentendo di mettere in relazione diretta le unità di un set con quelle dell'altro set dalle quali sono definite. Quindi, le classi di attori equivalenti, ossia le posizioni, vengono definite a partire dai loro modelli di partecipazione ad eventi, o appartenenza a gruppi, che, a loro volta, costituiscono classi di eventi, o gruppi, equivalenti in funzione dei modelli di partecipazione

degli attori che ne fanno parte. Questa informazione è molto utile per l'AFM in quanto consente di avere una visione sintetica della struttura del network direttamente sulla mappa, con indicazioni precise sulle relazioni di equivalenza che caratterizzano i cluster, ossia le posizioni, e di conseguenza il ruolo di network ricoperto all'interno della struttura relazionale. L'AFM, consentendo l'introdurre di informazioni aggiuntive offre la possibilità di integrare al suo interno, come variabili e punti supplementari, gli stessi risultati del Blockmodeling. Di conseguenza, il blockmodel può essere messo in relazione alla struttura dei legami sottostanti, e alle altre informazioni descrittive e strutturali, sia dal punto di vista analitico sia visuale, in un'ottica diacronica che tenga conto dell'evoluzione nel tempo. In effetti, uno dei vantaggi più interessanti è dato dal fatto che l'AFM, ma anche l'ACM, consente la proiezione delle posizioni in uno spazio metrico che rende possibile una rappresentazione pienamente significativa, in cui lo spazio tra le posizioni può essere interpretato in termini di distanze. Inoltre, l'opportunità di inserire le posizioni in un contesto informativo più ampio ne consente la caratterizzazione e una interpretazione più profonda. Infine, sembra importante sottolineare che la combinazione di queste tecniche comporta dei risultati interessanti anche da un punto di vista metodologico-concettuale. Le specificità tecniche dell'AFM consente di inserire in un approccio unitario analisi posizionale e teoria del campo. Tale approccio consente di combinare la dimensione della struttura relazionale di network con quella definita dalle caratteristiche delle unità che ne fanno parte e di cui si deve necessariamente tener conto per una piena comprensione dei fenomeni sociali che presentano un carattere complesso e multidimensionale.

Conclusioni

Il principale obiettivo di questo lavoro di tesi è stato quello di sviluppare una strategia di analisi orientata ad affrontare alcune delle problematiche legate all'applicazione delle tecniche di Blockmodeling per l'analisi dei network *two-mode*. Tale strategia ha previsto il ricorso all'ausilio degli strumenti offerti dalle tecniche di Analisi Fattoriale, in particolare all'utilizzo dell'Analisi delle Corrispondenze Multiple. Il primo passo nella formulazione della presente proposta è stato quello di definire i concetti di base legati alla prospettiva della Social Network Analysis, affrontati nel corso del primo capitolo. Il passo successivo è stato quello di circoscrivere il focus dell'analisi e individuare l'approccio più appropriato a perseguirlo. A tal proposito, l'interesse conoscitivo è stato rivolto all'analisi della struttura del network e all'individuazione, in esso, di sotto-strutture come entità rilevanti dotate di particolare significato. L'approccio adottato a tale scopo è stato quello posizionale che ha fornito parte degli strumenti concettuali e metodologici utilizzati. Nello specifico, l'approccio posizionale consente di svelare la struttura sottostante il network identificando in essa sottoinsiemi di unità e relazioni rilevanti, definite come posizioni e ruoli. Lo strumento privilegiato adottato per lo svolgimento di tale compito è stato il metodo del *two-mode* Blockmodeling, un'estensione del Blockmodeling Generalizzato per i network *two-mode* che consente specificamente di partizionare le unità di ognuno dei due set in cluster, ossia le posizioni, e le relazioni tra cluster in blocchi, a partire da una serie di tipi di blocchi ideali. Tali tipi blocchi ideali descrivono specifici tipi di equivalenza che riflettono specifici modelli di connessioni. Il Blockmodeling cerca di partizionare i legami tra i cluster in modo che i blocchi riflettano il più possibile i modelli di connessioni empirici. Dell'analisi di Blockmodeling e delle criticità ad essa legata si è ampiamente discusso nel secondo capitolo. Nel corso del terzo, invece, sono state

analizzate le relazioni tra il Blockmodeling e l'Analisi Fattoriale, in particolare l'Analisi delle Corrispondenze Multiple, che hanno suggerito l'uso combinato dei due metodi. Nello specifico, si è mostrato che questi ultimi sono particolarmente adatti ad analizzare i network *two-mode* in quanto ne consentono un approccio diretto, senza trasformazioni, rispettandone la natura duale e le caratteristiche specifiche dei suoi set che lo compongono. Inoltre, si è mostrato che il Blockmodeling e l'Analisi delle Corrispondenze Multiple, oltre ad essere unite dal comune intento di individuare e sintetizzare la struttura sottostante i fenomeni osservati, guardandoli da una prospettiva globale, sono legate dalla possibilità di rilevare delle proprietà simili delle unità. Infatti, entrambe consentono di ricostruire le strutture di similarità nel network sulla base di un certo criterio di equivalenza tra le unità. L'ACM si basa sul principio di equivalenza distribuzionale, mentre il Blockmodeling, almeno nella sua forma originaria, si basa sul principio di equivalenza strutturale. Queste due forme di equivalenza hanno una natura diversa. Nel primo caso è legata alle distribuzioni di frequenza dei profili di riga e di colonna, mentre, nel secondo caso alle relazioni tra le unità. Tuttavia, esiste una forte connessione tra questi due principi. Tale connessione diventa più forte ed evidente nel caso dell'utilizzo dell'ACM, poichè il complesso sistema di pesi che ne è alla base, legato alla metrica del chi-quadrato, rende le distanze che ne risultano molto simili alla distanza Euclidea ponderata. A sua volta, la distanza Euclidea è una misura sensibile e compatibile con l'equivalenza strutturale. Di conseguenza, le rappresentazioni grafiche realizzate attraverso l'ACM sono basate su proprietà che gli consentono di cogliere alcuni aspetti della struttura relazionale simili a quelli derivanti dal Blockmodeling basato sull'equivalenza strutturale. Su tali connessioni si basa la scelta di combinare le due tecniche.

L'utilità di questa combinazione, e in particolare dell'uso dell'ACM a supporto dell'analisi di Blockmodeling, che costituisce il cuore della presente proposta, è stata affrontata nel corso del quarto capitolo. Nello specifico si è mostrato che l'ACM può essere di supporto all'analisi di *two-mode* Blockmodeling lungo tre direzioni principali.

La prima direzione è stata sviluppata nell'ambito del *two-mode* Blockmodeling Esplorativo dove si è proposto di applicare il metodo di decomposizione di Eckart-Young (1936) ed eseguire il Blockmodeling su una matrice di affiliazione di rango

inferiore. Questa operazione ha consentito di estrarre la struttura relazionale forte, eliminando in parte le inconsistenze, con un vantaggio sia in termini computazionali sia nel trattamento di network ampie e/o sparse. Tale risultato è reso possibile dall'effetto di "regolarizzare i pattern relazionali" operato dal metodo di ricostruzione, che consente di eseguire una operazione di "filtraggio" rispetto ai legami che non sono coerenti con la struttura generale del network. In particolare la procedura tende a ricostruire pattern relazionali omogenei, "nulli" o "completi". In termini di Blockmodeling, questo significa eliminare legami presenti nei blocchi che dovrebbero essere nulli ed aggiungere legami nei blocchi che dovrebbero essere completi (considerata la struttura generale del network). L'operazione di ricostruzione basata sul metodo di Eckart-Young, quindi, risponde alla stessa logica alla base del Blockmodeling eseguito in base al criterio dell'equivalenza strutturale. La scelta di tale strategia come fase esclusiva è da valutare con cautela in quanto nelle struttura debole, che viene scartata insieme alle inconsistenze, potrebbero celarsi pattern relazionali di un certo interesse. Tale operazione di riduzione può essere suggerita come fase preliminare al Blockmodeling sul network originale e consente di ottenere una visione globale della struttura e di acquisire informazioni utili come guida per alcune scelte iniziali quali il numero di cluster. In quest'ottica, tale strategia è in linea con gli approcci di analisi statistica multidimensionale che mostrano come la capacità di recupero delle informazioni delle procedure di clustering possa essere migliorata, e resa meno sensibile al rumore presente nei dati, utilizzando i metodi fattoriali prima della *Cluster Analysis* (19, 138). Tale approccio non è esente da critiche (172), tuttavia, riteniamo che costituisca un primo passo verso la definizione di una strategia che preveda una vera e propria integrazione dei due metodi.

La scelta di tale strategia come fase esclusiva è da valutare con cautela in quanto nelle struttura debole, che viene scartata insieme alle inconsistenze, potrebbero celarsi pattern relazionali di un certo interesse. Tale operazione di riduzione può essere suggerita come fase preliminare al Blockmodeling sul network originale e consente di ottenere una visione globale della struttura e di acquisire informazioni utili come guida per alcune scelte iniziali quali il numero di cluster.

La seconda direzione è stata sviluppata nell'ambito del *two-mode* Blockmodeling Pre-specificato. L'ACM si è rivelata di supporto al *two-mode* Blockmodeling,

in quanto, insieme al numero di cluster, ha fornito una serie di informazioni utilizzate per l'individuazione dei tipi di blocchi ideali più appropriati al modello di connessioni; per la definizione di un modello di struttura del blockmodel ideale da testare; per la definizione dei vincoli da imporre alla struttura del clustering. L'ACM, consentendo l'introduzione di informazioni aggiuntive risulta molto efficace a tale scopo. I vantaggi di una strategia simile riguardano sia la riduzione del numero di tentativi necessari all'individuazione del numero di cluster e del blockmodel ottimale, sia la possibilità di avere una base di conoscenze sulla quale fondare la scelta tra partizioni egualmente ben modellate, con valori delle funzione criterio minimi. Inoltre, le informazioni sono state ottenute attraverso la ricostruzione dei pattern di connessioni e del contesto in cui le relazioni sono calate e dal quale possono essere influenzate. In questo modo, partizioni analoghe, rispetto ad un certo criterio di equivalenza, possono essere confrontate non solo in base a considerazioni di carattere teorico ma anche empirico. La logica di fondo dell'approccio da noi proposto si basa sul presupposto che la produzione della conoscenza segua un processo circolare guidato dai dati la cui esplorazione consentono l'individuazione di pattern di relazioni rilevanti a partire dai quali è possibile formulare delle ipotesi sulla struttura di dati da sottoporre a verifica attraverso un approccio confermativo che fornisce ulteriori intuizioni per formulare nuove ipotesi, ecc. Di conseguenza, nell'approccio di analisi vi è un'interdipendenza dell'aspetto induttivo e deduttivo, che parte dall'esplorazione dei pattern di relazioni attraverso l'AF che consente di definire un'ipotesi sul modello di struttura esistente nei dati da testare, e quindi verificare, attraverso l'analisi di Blockmodeling. Allo stesso tempo, i risultati del Blockmodeling consentono di riformulare e ri-specificare il modello di struttura più adeguato ai modelli di connessioni e i risultati che ne derivano possono essere ulteriormente introdotti in una cornice di analisi multidimensionale e messi in relazione al altre informazioni e migliorarne l'interpretazione.

Come terza direzione, e in modo non del tutto separato rispetto quelle precedenti, l'ACM ha offerto la possibilità di integrare al suo interno, come variabili e punti supplementari, gli stessi risulti del Blockmodeling. Di conseguenza, il blockmodel è stato messo in relazione alla struttura dei legami sottostanti, e alle altre informazioni descrittive e strutturali, sia dal punto di vista analitico sia

visuale. In effetti, uno dei vantaggi più interessanti è dato dal fatto che l'ACM consente la proiezione delle posizioni in uno spazio metrico che rende possibile una rappresentazione pienamente significativa, in cui lo spazio tra le posizioni può essere interpretato in termini di distanze. Un vantaggio ulteriore, in linea con quelli della direzione precedente, è stato dato dalla possibilità di ottenere informazioni analitiche utili ad orientare la scelta tra partizioni egualmente ben modellate (*equally well fitting partition*). In questo modo, partizioni analoghe, rispetto ad un certo criterio di equivalenza, possono essere confrontate non solo in base a considerazioni di carattere teorico ma anche empirico. Inoltre, l'opportunità di inserire le posizioni in un contesto informativo più ampio ne consente la caratterizzazione e una interpretazione più profonda.

Pertanto, l'integrazione di entrambe le tecniche ha offerto l'opportunità di valorizzare i singoli contributi informativi e migliorare l'interpretazione globale dei risultati con riflessi positivi sulla capacità di comprensione del fenomeno sociale in esame. Con tali premesse, la presente proposta può essere considerata come un approccio multi-step in cui le due tecniche possono interagire in una o più fasi del processo di analisi dei dati, a seconda delle specifiche esigenze.

Nel corso del quinto, e ultimo, capitolo, le opportunità derivanti dalla combinazione di tecniche di Analisi Fattoriale e tecniche di Blockmodeling sono state mostrate anche in una prospettiva più ampia legata alla gestione di network eterogenei. I fenomeni sociali sono complessi, multidimensionali, pertanto la loro comprensione richiede la capacità di mettere in relazione aspetti diversi a differenti livelli di analisi. Nel tentativo di compiere un passo in questa direzione è stata proposta una strategia di analisi che ha combinato l'Analisi Fattoriale Multipla (AFM) (87, 169) e il Blockmodeling per network *two-mode* eterogenei, composti da dati relativi ai legami tra due classi di entità differenti, in momenti temporali diversi, con attributi descritti delle entità e le loro posizioni strutturali derivanti dall'analisi di Blockmodeling. Tale approccio analitico è stato adottato per lo studio del campo delle co-produzioni teatrali nell'ambito del sistema teatrale regionale campano, in Italia (182). Nello specifico, lo studio ha analizzato la struttura del campo delle co-produzioni teatrali attraverso una integrazione della Teoria dei Campi di Pierre Bourdieu (35, 36) e dell'Approccio Posizionale (54, 56, 83) della SNA in una cornice di AFM. Pertanto, l'utilizzo combinato

dell'AFM e del Blockmodeling ha consentito unire la Teoria dei Campi e l'analisi Posizionale in un unico *framework* concettuale e metodologico per l'analisi dei network di affiliazione. Tale scelta ha portato ad un approccio unitario che ha consentito di tenere insieme l'analisi della dimensione relazionale, ossia la struttura relazionale, con quella legata alle proprietà descrittive delle unità, la struttura oggettiva. In questo modo è stato possibile cogliere gli elementi caratterizzanti della struttura delle co-produzione teatrali; individuare i pattern di connessioni principali e le loro caratteristiche; visualizzare le posizioni strutturali e caratterizzarle attraverso la loro collocazione nello spazio sociale definito dalle relazioni sottostanti e dagli attributi. Infine, è stato possibile osservare le variazioni di questi aspetti nel tempo. Come si è detto i risultati ottenuti sono relativi ai soli network analizzati. Pertanto, più che risultati definitivi relativi all'applicazione del metodo, quelli proposti, hanno lo scopo di evidenziare punti su cui le Tecniche dell'Analisi Fattoriale e quelle di Blockmodeling possono interagire, traendo beneficio dai rispettivi contributi. In quest'ottica, quindi, rappresentano dei suggerimenti alla base di ulteriori approfondimenti.

Tra i punti che richiedono ulteriore approfondimento possono essere considerati la necessità di individuare un criterio per stabilire il limite oltre il quale la procedura di ricostruzione legata al metodo di Eckart-Young deve essere interrotta per evitare che l'aggiunta di ulteriori informazioni genere una riconfigurazione dei legami con delle ricostruzioni errate. Inoltre, può essere interessante pensare alla creazione di un'unica funzione criterio che tenga insieme i vantaggi di riduzione dell'Analisi Fattoriale con quelli di partizionamento derivanti dal *two-mode* Blockmodeling, in direzione di una strategia simile a quella proposta da Vichi e Saporta (188).

Infine, sembra utile estendere tale approccio ad altre strutture di dati, come i network pesati. In questa direzione è stata da noi proposto l'utilizzo dell'Analisi delle Corrispondenze in prospettiva *Doubling* per network *two-mode* pesati, con valori dei legami espressi su scala ordinale (168). Utile sarebbe esplorare la relazione di questo approccio con quello del Blockmodeling per network pesati.

Appendice

Network *two-mode time-varying*

Un network *two-mode time-varying* può essere rappresentato attraverso un set di K network *two-mode* $\{\mathcal{G}_k\}_{k=1,\dots,K}$, in cui l'indice k si riferisce a differenti punti temporali in genere definite occasioni (*occasion*). Ognuno dei network *two-mode* \mathcal{G}_k è costituito da due set di unità (attori ed eventi) connessi da una relazione, e può essere rappresentato come in grafo bipartito $\mathcal{G}_k(\mathcal{V}_{1k}, \mathcal{V}_{2k}, R_k)$, composto da due set disgiunti di nodi - \mathcal{V}_{1k} e \mathcal{V}_{2k} rispettivamente, di cardinalità N_k e J_k - e un set di legami o archi, $R_k \subseteq \mathcal{V}_{1k} \times \mathcal{V}_{2k}$. Per definizione l'intersezione dei due set è un insieme vuoto, $\mathcal{V}_{1k} \cap \mathcal{V}_{2l} = \emptyset, \forall k$.

Nella matrice considerata il primo set di attori è fisso nel tempo $\mathcal{V}_{1k} = \mathcal{V}_1 \forall k$, mentre il secondo può variare. Pertanto il set $\mathcal{V}_1 = \{a_1, a_2, \dots, a_N\}$ rappresenta il set di N attori, mentre $\mathcal{V}_{2k} = \{e_{1k}, e_{2k}, \dots, e_{J_k}\}$ rappresenta il set di J_k eventi. Il legame $r_{ijk} = (a_i, e_{jk})$, $r_{ijk} \in R_k$ è una coppia ordinata di nodi e indica se un attore a_i partecipa ad un evento e_{jk} , o meno.

Ogni set $\mathcal{V}_1 \times \mathcal{V}_{2k}$ può essere pienamente rappresentato attraverso una matrice di affiliazione $\mathbf{F}_k = (f_{ijk})$, $i = 1, \dots, N$, $j = 1, \dots, J_k$, $k = 1, \dots, K$, con $f_{ijk} = 1$ se $(a_i, e_{jk}) \in R_k$ e 0 altrimenti. Data \mathbf{F}_k , i marginali di riga e di colonna $f_{i.k} = \sum_{j=1}^{J_k} f_{ijk}$ e $f_{.jk} = \sum_{i=1}^N f_{ijk}$ corrispondono, rispettivamente, al *degree* d_{ik} dell' i -esimo attore nell'occasione k e alla *size* s_{jk} del j -esimo evento nell'occasione k , quindi, $f_{i.k} = d_{ik}$ e $f_{.jk} = s_{jk}$. Il set di tutte le matrici di affiliazione $\{\mathbf{F}_k\}_{k=1,\dots,K}$ possono essere rappresentati attraverso una tabella di dati, in genere chiamata *grand table*, $\mathbb{F} = [\mathbf{F}_1 | \dots | \mathbf{F}_K] = (f_{ijk})$, con $f_{ijk} \in \{0, 1\}$, $i = 1, \dots, N$, $j = 1, \dots, J_k$, $k = 1, \dots, K$. La *grand table* \mathbb{F} è costruita attraverso l'affiancamento di sotto-tabelle $\{\mathbf{F}_k\}_{k=1,\dots,K}$ (fig.5.18). Data \mathbb{F} , per ogni i , $i = 1, \dots, N$, il marginale

5.5 Analisi Fattoriale Multipla e Blockmodeling: considerazioni finali

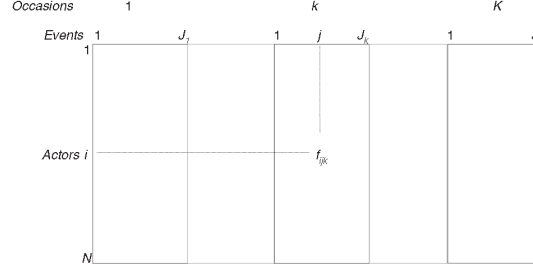


Figura 5.18: *Grand table* \mathbb{F} per un network *two-mode time varying*. L'elemento f_{ijk} della matrice di affiliazione \mathbf{F}_k codifica la partecipazione/non-partecipazione dell'attore i all'evento j per l'occasione k .

di riga $f_{i..} = \sum_{j=1}^{J_k} \sum_{k=1}^K f_{ijk}$ coincide con il *degree* totale dell' i -esimo attore, $d_i = d_{i..}$, e $L = \sum_{i=1}^N \sum_{j=1}^{J_k} \sum_{k=1}^K f_{ijk}$ è il numero totale di legami su tutte le occasioni ¹.

L'applicazione dell'AFM per l'analisi dei network *two-mode time varying*, che includono come variabili supplementari anche le posizioni degli attori/eventi, consente di: (i) analizzare la struttura delle relazioni incorporata in ogni matrice di affiliazione \mathbf{F}_K in termini di similarità relazionale tra attori/eventi, (*analisi parziale*); (ii) analizzare in modo cumulativo le occasioni K , cercando la struttura relazionale complessiva racchiusa nella matrice di affiliazione \mathbb{F} , (considerando anche la collocazione delle posizioni degli eventi/attori in relazione agli attori/eventi a cui sono associati e rispetto alle loro caratteristiche descrittive) (*analisi globale*); (iii) evidenziare le differenze tra le occasioni K in termini di struttura globale (variazioni strutturali tra le occasioni, osservando i cambiamenti nel pattern relazionali sia nel tempo sia nei due modi); (iv) analizzare le variazioni degli attori/eventi nelle varie occasioni (osservando le variazioni nei pattern relazionali proiettando le matrici di affiliazione pesate \mathbf{F}_k sul piano fattoriale globale). Gli ultimi due punti, (iii) e (iv), permettono di studiare il cambiamento longitudinale e il comportamento nel tempo dei modelli relazionali degli attori/eventi e delle posizioni a cui sono associati.

¹Per una trattazione dettagliata sull'applicazione del metodo dell'AFM ai network *two-mode time-varying* si rinvia a Ragozini et al. (169).

Bibliografia

- [1] AGNEESSENS, F. & EVERETT, M.G. (2013). Introduction to the special issue on advances in two-mode social networks. *Social Networks*, **2**, 145–147.
- [2] ALNOUKARI, M. & EL SHEIKH, A. (2011). Knowledge discovery process models: From traditional to agile modeling. *Business Intelligence and Agile Methodologies for Knowledge-Based Organizations: Cross-Disciplinary Applications*, 72–100.
- [3] ANDERSON, C.J., WASSERMAN, S. & FAUST, K. (1992). Building stochastic blockmodels. *Social networks*, **14**, 137–161.
- [4] ANHEIER, H.K., GERHARDS, J. & ROMO, F.P. (1995). Forms of capital and social structure in cultural fields: Examining bourdieu’s social topography. *American journal of sociology*, 859–903.
- [5] BARNES, J.A. (1954). *Class and committees in a Norwegian island parish*. Plenum New York. 2
- [6] BARNETT, G.A. & DANOWSKI, J.A. (1992). The structure of communication. *Human Communication Research*, **19**, 264–285. 85
- [7] BATAGELJ, V. (1997). Notes on blockmodeling. *Social Networks*, **19**, 143–155. 50, 52, 54
- [8] BATAGELJ, V. (2011). Large-scale network analysis. In SCOTT, J., & CARRINGTON, P., (eds.), *The SAGE handbook of social network analysis*, 550–557, SAGE publications.

- [9] BATAGELJ, V. & BREN, M. (1995). Comparing resemblance measures. *Journal of classification*, **12**, 73–90. 38
- [10] BATAGELJ, V. & FERLIGOJ, A. (2000). *Clustering relational data*. Springer. 48, 56, 63, 65
- [11] BATAGELJ, V. & MRVAR, A. (1998). Pajek-program for large network analysis. *Connections*, **21**, 47–57. 69, 127, 138
- [12] BATAGELJ, V. & MRVAR, A. (2001). Pajek analysis and visualization of large networks. In *Graph Drawing*, 477–478, Springer. 106
- [13] BATAGELJ, V., DOREIAN, P. & FERLIGOJ, A. (1992). An optimizational approach to regular equivalence. *Social Networks*, **14**, 121–135. 52, 54, 65
- [14] BATAGELJ, V., FERLIGOJ, A. & DOREIAN, P. (1992). Direct and indirect methods for structural equivalence. *Social networks*, **14**, 63–90.
- [15] BATAGELJ, V., FERLIGOJ, A. & DOREIAN, P. (1998). *Fitting pre-specified blockmodels*. Springer. 71
- [16] BATAGELJ, V., MRVAR, A., FERLIGOJ, A. & DOREIAN, P. (2004). Generalized blockmodeling with pajek. *Metodoloski zvezki*, **1**, 455–467. 47, 48
- [17] BATAGELJ, V., FERLIGOJ, A. & DOREIAN, P. (2007). *Indirect blockmodeling of 3-way networks*. Springer. 71
- [18] BATAGELJ, V., DOREIAN, P., FERLIGOJ, A. & KEJZAR, N. (2014). *Understanding large temporal networks and spatial networks: Exploration, pattern searching, visualization and network evolution*. John Wiley & Sons. 106
- [19] BÉCUE-BERTAUT, M. & PAGÈS, J. (2008). Multiple factor analysis and clustering of a mixture of quantitative, categorical and frequency data. *Computational Statistics & Data Analysis*, **52**, 3255–3268.

- [20] BENZÉCRI, J.P. *et al.* (1973). *L'Analyse des Données. Tome II: L'analyse des correspondances*. Dunod: Paris.
- [21] BERGE, C. & MINIEKA, E. (1973). *Graphs and hypergraphs*, vol. 7. North-Holland publishing company Amsterdam. 33
- [22] BERKOWITZ, L. (1982). *An Introduction to Structural Analysis, The Network Approach to Social Research*. Butterworth. 3
- [23] BOORMAN, S.A. & WHITE, H.C. (1976). Social structure from multiple networks. ii. role structures. *American journal of sociology*, 1384–1446. 3, 46, 47, 48
- [24] BORGATTI, S.P. (2009). Social network analysis, two-mode concepts in. In MEYERS, R. A., (eds.), *Encyclopedia of Complexity and Systems Science*, 8279–8291, Springer.
- [25] BORGATTI, S.P. & EVERETT, M.G. (1989). The class of all regular equivalences: Algebraic structure and computation. *Social Networks*, **11**, 65–88.
- [26] BORGATTI, S.P. & EVERETT, M.G. (1992). Regular blockmodels of multiway, multimode matrices. *Social Networks*, **14**, 91–120. 45, 48
- [27] BORGATTI, S.P. & EVERETT, M.G. (1997). Network analysis of 2-mode data. *Social networks*, **19**, 243–269. 85
- [28] BORGATTI, S.P. & EVERETT, M.G. (1992). Notions of position in social network analysis. *Sociological methodology*, **22**, 1–35. 46, 47, 48, 71, 72
- [29] BORGATTI, S.P. & FOSTER, P.C. (2003). The network paradigm in organizational research: A review and typology. *Journal of management*, **29**, 991–1013. 8
- [30] BORGATTI, S.P. & HALGIN, D.S. (2011). Analyzing affiliation networks. In SCOTT, J., & CARRINGTON, P., (eds.), *The SAGE handbook of social network analysis*, 417–433, SAGE publications.

- [31] BORGATTI, S.P., MEHRA, A., BRASS, D.J. & LABIANCA, G. (2009). Network analysis in the social sciences. *science*, **323**, 892–895. 7
- [32] BORGATTI, S.P., BRASS, D.J. & HALGIN, D.S. (2014). Social network research: Confusions, criticisms, and controversies. *Research in the sociology of organizations*, **40**, 1–29. 6, 7
- [33] BOTT, E. (1955). Urban families: conjugal roles and social networks. *Human relations*. 2
- [34] BOTTERO, W. & CROSSLEY, N. (2011). Worlds, fields and networks: Becker, bourdieu and the structures of social relations. *Cultural Sociology*, **5**, 99–119.
- [35] BOURDIEU, P. (1983). The field of cultural production, or: The economic world reversed. *Poetics*, **12**, 311–356.
- [36] BOURDIEU, P. (1984). *Distinction: A social critique of the judgement of taste*. Harvard University Press.
- [37] BOURDIEU, P. (1988). *Homo academicus*. Stanford University Press.
- [38] BOURDIEU, P. (1989). Social space and symbolic power. *Sociological theory*, **7**, 14–25.
- [39] BOURDIEU, P. (2010). The forms of capital. (1986). In SZEMAN, I., & KAPOSY, T., (eds.), *Cultural theory: an anthology*, 81–93, John Wiley & Sons.
- [40] BOURDIEU, P. & JOHNSON, R. (1993). *The field of cultural production: Essays on art and literature*. Columbia University Press.
- [41] BOURDIEU, P. & WACQUANT, L.J. (1992). *An invitation to reflexive sociology*. University of Chicago press.
- [42] BRANDES, U., FLEISCHER, D. & PUPPE, T. (2007). Dynamic spectral layout with an application to small worlds. *J. Graph Algorithms Appl.*, **11**, 325–343. 138

- [43] BREIGER, R.L. (1974). The duality of persons and groups. *Social forces*, **53**, 181–190. 32, 44, 50, 78, 84, 85
- [44] BREIGER, R.L. (2000). A tool kit for practice theory. *Poetics*, **27**, 91–115. 86, 96
- [45] BREIGER, R.L., BOORMAN, S.A. & ARABIE, P. (1975). An algorithm for clustering relational data with applications to social network analysis and comparison with multidimensional scaling. *Journal of mathematical psychology*, **12**, 328–383.
- [46] BRUSCO, M. & DOREIAN, P. (2015). A real-coded genetic algorithm for two-mode kl-means partitioning with application to homogeneity blockmodeling. *Social Networks*, **41**, 26–35. 72, 75
- [47] BRUSCO, M. & STEINLEY, D. (2006). Inducing a blockmodel structure of two-mode binary data using seriation procedures. *Journal of Mathematical Psychology*, **50**, 468–477. 71, 75
- [48] BRUSCO, M. & STEINLEY, D. (2007). A variable neighborhood search method for generalized blockmodeling of two-mode binary matrices. *Journal of Mathematical Psychology*, **51**, 325–338. 75
- [49] BRUSCO, M. & STEINLEY, D. (2011). A tabu-search heuristic for deterministic two-mode blockmodeling of binary network matrices. *Psychometrika*, **76**, 612–633.
- [50] BRUSCO, M., DOREIAN, P., LLOYD, P. & STEINLEY, D. (2013). A variable neighborhood search method for a two-mode blockmodeling problem in social network analysis. *Network Science*, **1**, 191–212. 63, 71, 75
- [51] BRUSCO, M., DOREIAN, P., MRVAR, A. & STEINLEY, D. (2013). An exact algorithm for blockmodeling of two-mode network data. *The Journal of Mathematical Sociology*, **37**, 61–84. 71, 75
- [52] BRUSCO, M., DOREIAN, P., STEINLEY, D. & SATORNINO, C.B. (2013). Multiobjective blockmodeling for social network analysis. *Psychometrika*, **78**, 498–525. 71, 75

- [53] BURT, R. & MINOR, M. (1983). *Applied networks analysis*. Berverly Hills, Sage.
- [54] BURT, R.S. (1976). Positions in networks. *Social forces*, **55**, 93–122. 47, 50, 81, 103, 137
- [55] BURT, R.S. (1978). A structural theory of interlocking corporate directorates. *Social Networks*, **1**, 415–435.
- [56] BURT, R.S. (1980). Models of network structure. *Annual review of sociology*, 79–141. 8, 39, 40, 41, 46, 81, 84, 97, 103, 137
- [57] CARROLL, J.D., GREEN, P.E. & SCHAFFER, C.M. (1986). Interpoint distance comparisons in correspondence analysis. *Journal of Marketing Research*, 271–280.
- [58] CARROLL, J.D., GREEN, P.E. & SCHAFFER, C.M. (1987). Comparing interpoint distances in correspondence analysis: A clarification. *Journal of Marketing Research*, 445–450.
- [59] CARROLL, J.D., GREEN, P.E. & SCHAFFER, C.M. (1989). Reply to greenacre’s commentary on the carroll-green-schaffer scaling of two-way correspondence analysis solutions. *Journal of Marketing Research*, 366–368.
- [60] CARTWRIGHT, D. & HARARY, F. (1956). Structural balance: a generalization of heider’s theory. *Psychological review*, **63**, 277. 2
- [61] CASSE, J.I., SHELTON, C.R. & HANNEMAN, R.A. (2013). A new criterion function for exploratory blockmodeling for structural and regular equivalence. *Social Networks*, **35**, 31–50. 62
- [62] CHIESI, A.M. (1999). *L’analisi dei reticoli*, vol. 11. FrancoAngeli. 11, 18, 19, 20, 21, 25, 26, 46
- [63] CHUNG, C.J. & PARK, H.W. (2010). Textual analysis of a political message: The inaugural addresses of two korean presidents. *Social science information*, **49**, 215–239. 84

- [64] COX, T.F. & COX, M.A. (2000). *Multidimensional scaling*. CRC press. 44
- [65] CROSSLEY, N. (2008). Pretty connected the social network of the early uk punk movement. *Theory, Culture & Society*, **25**, 89–116.
- [66] DAVIS, A., GARDNER, B.B. & GARDNER, M.R. (1941). Deep south: A social anthropological study of class and caste.
- [67] DE NOOY, W. (2002). The dynamics of artistic prestige. *Poetics*, **30**, 147–167.
- [68] DE NOOY, W. (2003). Fields and networks: correspondence analysis and social network analysis in the framework of field theory. *Poetics*, **31**, 305–327.
- [69] DEGENNE, A. & FORSÉ, M. (1999). *Introducing social networks*. Sage. 4, 40
- [70] D’ESPOSITO, M.R., DE STEFANO, D. & RAGOZINI, G. (2014). A comparison of *Chi-square* metrics for the assessment of relational similarities in affiliation networks. In VICARI, D., OKADA, A., RAGOZINI, G., & WEIHS, C., (eds.), *Analysis and Modeling of Complex Data in Behavioral and Social Sciences, Studies in Classification, Data Analysis, and Knowledge Organization*, 113–122, Springer.
- [71] D’ESPOSITO, M.R., DE STEFANO, D. & RAGOZINI, G. (2014). On the use of multiple correspondence analysis to visually explore affiliation networks. *Social Networks*, **38**, 28–40.
- [72] DEZA, M.M. & DEZA, E. (2009). *Encyclopedia of distances*. Springer Berlin Heidelberg. 35
- [73] DIMAGGIO, P. (1986). Structural analysis of organizational fields: A blockmodel approach. *Research in organizational behavior*.
- [74] DIMAGGIO, P. (2011). Cultural networks. In SCOTT, J., & CARRINGTON, P., (eds.), *The SAGE handbook of social network analysis*, 286–310, SAGE publications.

-
- [75] DOREIAN, P. (2006). Some open problem sets for generalized blockmodeling. In BATAGELJ, V., FERLIGOJ, A., & ŽIBERNA, A., (eds.), *Data Science and Classification*, 119–130, Springer. 76
- [76] DOREIAN, P. (2008). A multiple indicator approach to blockmodeling signed networks. *Social Networks*, **30**, 247–258.
- [77] DOREIAN, P. (2009). Positional analysis and blockmodeling. In MEYERS, R. A., (eds.), *Encyclopedia of Complexity and Systems Science*, 6913–6927, Springer, Springer.
- [78] DOREIAN, P. & MRVAR, A. (1996). A partitioning approach to structural balance. *Social networks*, **18**, 149–168. 72
- [79] DOREIAN, P. & MRVAR, A. (2009). Partitioning signed social networks. *Social Networks*, **31**, 1–11. 47, 48, 50, 53, 72, 75, 76, 77
- [80] DOREIAN, P., BATAGELJ, V. & FERLIGOJ, A. (1994). Partitioning networks based on generalized concepts of equivalence. *Journal of Mathematical Sociology*, **19**, 1–27.
- [81] DOREIAN, P., BATAGELJ, V. & FERLIGOJ, A. (2004). Generalized blockmodeling of two-mode network data. *Social networks*, **26**, 29–53. 60, 71, 72, 73, 74, 77, 78
- [82] DOREIAN, P., BATAGELJ, V. & FERLIGOJ, A. (2005). Positional analyses of sociometric data. In CARRINGTON, P., SCOTT, J., & WASSERMAN, S., (eds.), *Models and methods in social network analysis*, 77–96, NY: Cambridge University Press. 48, 49, 50, 51, 52, 53, 54, 60, 63, 64, 65, 67
- [83] DOREIAN, P., BATAGELJ, V. & FERLIGOJ, A. (2005 a). *Generalized blockmodeling*, vol. 25. Cambridge university press. 38, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 77, 78, 81, 106, 123, 127, 137

- [84] DOREIAN, P., LLOYD, P. & MRVAR, A. (2013). Partitioning large signed two-mode networks: Problems and prospects. *Social Networks*, **35**, 178–203.
- [85] ECKART, C. & YOUNG, G. (1936). The approximation of one matrix by another of lower rank. *Psychometrika*, **1**, 211–218. 79, 107
- [86] EMIRBAYER, M. & JOHNSON, V. (2008). Bourdieu and organizational analysis. *Theory and society*, **37**, 1–44.
- [87] ESCOPIER, B. & PAGÈS, J. (1988). *Analyses factorielles simples et multiples: objectifs, méthodes et interprétation*. Dunod. 138
- [88] ESCOPIER, B. & PAGES, J. (1994). Multiple factor analysis (afmult package). *Computational statistics & data analysis*, **18**, 121–140.
- [89] EVERETT, M.G. & BORGATTI, S.P. (1994). Regular equivalence: General theory. *Journal of mathematical sociology*, **19**, 29–52. 49
- [90] EVERETT, M.G. & BORGATTI, S.P. (1999). The centrality of groups and classes. *The Journal of mathematical sociology*, **23**, 181–201. 53, 71
- [91] EVERETT, M.G. & BORGATTI, S.P. (2013). The dual-projection approach for two-mode networks. *Social Networks*, **35**, 204–210. 71
- [92] FAUST, K. (1988). Comparison of methods for positional analysis: Structural and general equivalences. *Social Networks*, **10**, 313–341. 97
- [93] FAUST, K. (1997). Centrality in affiliation networks. *Social networks*, **19**, 157–191.
- [94] FAUST, K. (2005). Using correspondence analysis for joint displays of affiliation networks. In CARRINGTON, P., SCOTT, J., & WASSERMAN, S., (eds.), *Models and methods in social network analysis*, 117–147, Cambridge University Press, Cambridge. 44, 86

- [95] FAUST, K. & ROMNEY, A.K. (1985). Does structure find structure?: A critique of burt's use of distance as a measure of structural equivalence. *Social Networks*, **7**, 77–103. 97
- [96] FAUST, K. & WASSERMAN, S. (1992). Blockmodels: Interpretation and evaluation. *Social networks*, **14**, 5–61. 46, 47, 56
- [97] FELD, S.L. (1981). The focused organization of social ties. *American journal of sociology*, 1015–1035. 85
- [98] FERLIGOJ, A., DOREIAN, P. & BATAGELJ, V. (2011). Positions and roles. In SCOTT, J., & CARRINGTON, P., (eds.), *The SAGE handbook of social network analysis*, 434–446, SAGE publications. 76, 106
- [99] FIELD, S., FRANK, K.A., SCHILLER, K., RIEGLE-CRUMB, C. & MULLER, C. (2006). Identifying positions from affiliation networks: Preserving the duality of people and events. *Social Networks*, **28**, 97–123. 85
- [100] FORTUNATO, S. (2010). Community detection in graphs. *Physics reports*, **486**, 75–174. 45
- [101] FOSTER, B.L. & SEIDMAN, S.B. (1982). Urban structures derived from collections of overlapping subsets. *Urban anthropology*, 177–192. 32
- [102] FREEMAN, L. (2004). *The development of social network analysis*. Empirical Press Vancouver, BC Canada. 2, 3, 5, 40
- [103] FREEMAN, L.C. (1993). Finding groups with a simple genetic algorithm*. *Journal of Mathematical Sociology*, **17**, 227–241. 53
- [104] FREEMAN, L.C. (2000). Visualizing social networks. *Journal of social structure*, **1**, 4. 138
- [105] FREEMAN, L.C. (2005). Graphic techniques for exploring social network data. In CARRINGTON, P., SCOTT, J., & WASSERMAN, S., (eds.), *Models and methods in social network analysis*, 248–269, NY: Cambridge University Press. 40, 44, 133

- [106] FREEMAN, L.C. (2011). The development of social network analysis-with an emphasis on recent events. In SCOTT, J., & CARRINGTON, P., (eds.), *The SAGE handbook of social network analysis*, 26–39, SAGE publications. 5, 39
- [107] FRIEDKIN, N.E. (1984). Structural cohesion and equivalence explanations of social homogeneity. *Sociological Methods & Research*, **12**, 235–261. 41
- [108] GABRIEL, K. (1995). Biplot display of multivariate categorical data, with comments on multiple correspondence analysis. In KRZANOWSKY, W., (eds.), *Recent advances in descriptive multivariate analysis*, 190–226, Oxford, UK: Clarendon Press.
- [109] GALASKIEWICZ, J., WASSERMAN, S., RAUSCHENBACH, B., BIELEFELD, W. & MULLANEY, P. (1985). The influence of corporate power, social status, and market position on corporate interlocks in a regional network. *Social Forces*, **64**, 403–431.
- [110] GERHARDS, J. & ANHEIER, H.K. (1989). The literary field: An empirical investigation of bourdieu’s sociology of art. *International Sociology*, **4**, 131–146. 96
- [111] GIORDANO, G. & VITALE, M.P. (2006). Multidimensional data analysis to assess interactions in an e-learning community. *Journal of e-Learning and Knowledge Society*, **2**, 191–204.
- [112] GIORDANO, G. & VITALE, M.P. (2007). Factorial contiguity maps to explore relational data patterns. *Statistica Applicata*, **19**, 297–306. 44, 45
- [113] GIORDANO, G. & VITALE, M.P. (2009). On the use of multidimensional data analysis techniques for social networks. In AA.VV. (eds.), *Statistical Methods for the analysis of large data-sets, Italian Statistical Society. Book of short papers*, 93–96, CLEUP Editore.
- [114] GIUFFRÉ, K. (1999). Sandpiles of opportunity: success in the art world. *Social forces*, **77**, 815–832.

- [115] GOODMAN, R.A. & GOODMAN, L.P. (1976). Some management issues in temporary systems: A study of professional development and manpower-the theater case. *Administrative Science Quarterly*, 494–501.
- [116] GOWER, J. (2006). Divided by a common language: Analyzing and visualizing two-way arrays. In GREENACRE, M., & BLASIUS, J., (eds.), *Multiple correspondence analysis and related methods*, 77–105, Chapman and Hall CRC Press.
- [117] GOWER, J. (2014). Some history of algebraic canonical forms and data analysis. In BLASIUS, J., & GREENACRE, M., (eds.), *Visualization and Verbalization of Data*, 17–30, Chapman and Hall CRC Press.
- [118] GREENACRE, M. & BLASIUS, J. (2006). *Multiple correspondence analysis and related methods*. CRC Press. 86
- [119] GREENACRE, M.J. (1984). *Theory and applications of correspondence analysis*. Academic Press. 79, 86, 95
- [120] GREENACRE, M.J. (1989). The carroll-green-schaffer scaling in correspondence analysis: a theoretical and empirical appraisal. *Journal of Marketing Research*, 358–365.
- [121] GREENACRE, M.J. (2010). *Biplots in practice*. Fundacion BBVA.
- [122] HANDCOCK, M.S., RAFTERY, A.E. & TANTRUM, J.M. (2007). Model-based clustering for social networks. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, **170**, 301–354.
- [123] HANNEMAN, R.A. & RIDDLE, M. (2005). Introduction to social network methods. 40, 43
- [124] HANNEMAN, R.A. & RIDDLE, M. (2011). Concepts and measures for basic network analysis. In SCOTT, J., & CARRINGTON, P., (eds.), *The SAGE handbook of social network analysis*, 340–369, SAGE publications. 7, 17, 30, 33, 40
- [125] HARARY, F. (1969). *Graph theory*. Reading, MA: Addison-Wesley. 18

- [126] HARARY, F. & NORMAN, R.Z. (1953). *Graph theory as a mathematical model in social science*. University of Michigan Press. 10
- [127] HOLLAND, P.W., LASKEY, K.B. & LEINHARDT, S. (1983). Stochastic blockmodels: First steps. *Social networks*, **5**, 109–137.
- [128] HORVÁT, E.A. & ZWEIG, K.A. (2012). One-mode projection of multiplex bipartite graphs. In *Proceedings of the 2012 International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2012)*, 599–606, IEEE Computer Society.
- [129] HSIEH, M.H. & MAGEE, C.L. (2008). An algorithm and metric for network decomposition from similarity matrices: Application to positional analysis. *Social Networks*, **30**, 146–158. 106
- [130] JACCARD, P. (1908). Nouvelles recherches sur la distribution florale. *Bulletin de la Societe Vaudoise des Sciences Naturelles*, **44**, 223–270.
- [131] KNOKE, D. & KUKLINSKI, J.H. (1982). *Network analysis*.. Beverly Hills Calif. Sage Publications. 9, 48
- [132] KOLACZYK, E. (2009). *Statistical Analysis of Network Data, volume 69 of Springer Series in Statistics*. Springer New York. 4
- [133] KOVÁCS, B. (2010). A generalized model of relational similarity. *Social Networks*, **32**, 197–211. 97
- [134] KROONENBERG, P.M. (2008). *Applied multiway data analysis*, vol. 702. John Wiley & Sons. 138
- [135] LAUMANN, E.O. (1966). *Prestige and Association in an Urban Community*. Prentice-Hall. 3
- [136] LAURO, N.C. & GHERGHI, M. (2010). *Appunti di analisi dei dati multidimensionali: metodologie ed esempi*. RCE Multimedia Napoli. 109

- [137] LE ROUX, B. & ROUANET, H. (2004). *Geometric data analysis: from correspondence analysis to structured data analysis*. Springer Science & Business Media.
- [138] LEBART, L. (1994). Complementary use of correspondence analysis and cluster analysis. In GREENACRE, M., & BLASIUS, J., (eds.), *Correspondence Analysis in the Social Sciences*, 162–178, Academic Press, San Diego.
- [139] LEE, S. (2008). *Understanding the Nature of Citation: Examination of Structure in Citation Analysis and Its Application to Communication Discipline in the Information Age*. ProQuest. 85
- [140] LEWIN, K. (1936). *Principles of Topological Psychology*. New York: McGraw Hill. 1, 2
- [141] LINTON, R. (1936). *The study of man: an introduction*. Appleton-Century. 46
- [142] LORRAIN, F. & WHITE, H.C. (1971). Structural equivalence of individuals in social networks. *The Journal of mathematical sociology*, **1**, 49–80. 47, 97
- [143] MARIN, A. & WELLMAN, B. (2011). Social network analysis: An introduction. In SCOTT, J., & CARRINGTON, P., (eds.), *The SAGE handbook of social network analysis*, 11–25, Sage London. 7
- [144] MARSDEN, P.V. (1990). Network data and measurement. *Annual review of sociology*, 435–463. 6, 7
- [145] MARSDEN, P.V. (2011). Survey methods for network data. In SCOTT, J., & CARRINGTON, P., (eds.), *The SAGE handbook of social network analysis*, 370–388, SAGE publications.
- [146] MAYO, E. (1933). *The human problems of an industrial civilization*. The Macmillan Company.
- [147] MCANDREW, S. & EVERETT, M. (2015). Music as collective invention: A social network analysis of composers. *Cultural Sociology*, **9**, 56–80.

- [148] McGRATH, C., BLYTHE, J. & KRACKHARDT, D. (1997). The effect of spatial arrangement on judgments and errors in interpreting graphs. *Social Networks*, **19**, 223–242.
- [149] MEMON, B.R. & WIL, U.K. (2013). Visual analysis of heterogeneous networks. In *Intelligence and Security Informatics Conference (EISIC), 2013 European*, 129–134, IEEE. 136
- [150] MERTON, R.K. (1957). The role-set: Problems in sociological theory. *The British Journal of Sociology*, **8**, 106–120. 46
- [151] MICHAELSON, A. & CONTRACTOR, N.S. (1992). Structural position and perceived similarity. *Social Psychology Quarterly*, 300–310. 45
- [152] MILGRAM, S. (1967). The "small world" problem. *Psychology today*, **2**, 60–67. 4
- [153] MITCHELL, J.C. (1969). *Social networks in urban situations: analyses of personal relationships in Central African towns*. Manchester University Press.
- [154] MOODY, J. (2004). The structure of a social science collaboration network: Disciplinary cohesion from 1963 to 1999. *American sociological review*, **69**, 213–238. 84
- [155] MORENO, J.L. (1934). *Who shall survive*, vol. 58. JSTOR. 1
- [156] MRVAR, A. & BATAGELJ, V. (2004). Relinking marriages in genealogies. *Metodološki zvezki-Advances in Methodology and Statistics*, **1**, 407–418. 106
- [157] NADEL, S.F. (1957). *The Theory of Social Structure*. London: Cohen and West. 46
- [158] NENADIC, O. & GREENACRE, M. (2007). Correspondence analysis in r, with two-and three-dimensional graphics: The ca package. *Journal of Statistical Software*, **20**, 1–13.

- [159] NEWMAN, M.E. (2004). Detecting community structure in networks. *The European Physical Journal B-Condensed Matter and Complex Systems*, **38**, 321–330. 40, 45
- [160] NEWMAN, M.E. (2006). Finding community structure in networks using the eigenvectors of matrices. *Physical review E*, **74**, 036104. 43, 45
- [161] NEWMAN, M.E. & GIRVAN, M. (2004). Finding and evaluating community structure in networks. *Physical review E*, **69**, 026113. 45
- [162] OBERT R. SOKAL & SNEATH, P.H.A. (1963). *Principles of numerical taxonomy*. WH Freeman.
- [163] PAGÈS, J. (2002). Analyse factorielle multiple appliquée aux variables qualitatives et aux données mixtes. *Revue de statistique appliquée*, **50**, 5–37.
- [164] PAGÈS, J. (2014). *Multiple Factor Analysis by example using R*. CRC Press.
- [165] PISELLI, F. (2001). *Reti: l'analisi di network nelle scienze sociali*. Donzelli editore. 2, 3
- [166] PIZARRO, N. (2007). Structural identity and equivalence of individuals in social networks beyond duality. *International Sociology*, **22**, 767–792.
- [167] PROTA, L. & DOREIAN, P. (2016). Finding roles in sparse economic hierarchies: Going beyond regular equivalence. *Social Networks*, **45**, 1–17. 72, 112
- [168] RAGOZINI, G., DE STEFANO, D. & D'AMBROSIO, D. (2011). Correspondence analysis with doubling for two-mode valued networks. In MOLA, F., & CONVERSANO, C., (eds.), *Book of abstracts, CLADAG 2015 - 10th Scientific Meeting of the Classification and Data Analysis Group. Cagliari, October 8-10, 2015, Cagliari*, 343–346, CUEC Editrice.

- [169] RAGOZINI, G., DE STEFANO, D. & D'ESPOSITO, M.R. (2015). Multiple factor analysis for time-varying two-mode networks. *Network Science*, **3**, 18–36. 45, 81, 137, 138
- [170] REICHARDT, J. & WHITE, D.R. (2007). Role models for complex networks. *The European Physical Journal B*, **60**, 217–224. 46, 62, 106
- [171] ROBERTS, J.M. (2000). Correspondence analysis of two-mode network data. *Social Networks*, **22**, 65–72. 44, 86, 97
- [172] ROCCI, R., GATTONE, S.A. & VICHI, M. (2011). A new dimension reduction method: Factor discriminant k-means. *Journal of classification*, **28**, 210–226.
- [173] ROFFILLI, M. & LOMI, A. (2006). Identifying and classifying social groups: A machine learning approach. In BATAGELJ, V., FERLIGOJ, A., & ŽILBERNA, A., (eds.), *Data Science and Classification*, 149–157, Springer. 71
- [174] RUSSELL, P.F., RAO, T.R. *et al.* (1940). On habitat and association of species of anopheline larvae in south-eastern madras. *Journal of the Malaria Institute of India*, **3**, 153–178.
- [175] SAILER, L.D. (1978). Structural equivalence: Meaning and definition, computation and application. *Social Networks*, **1**, 73–90.
- [176] SCOTT, J. (2000). *Social network analysis*. Sage. 1, 5, 30, 33, 44
- [177] SCOTT, J. (2010). Social network analysis: developments, advances, and prospects. *Social network analysis and mining*, **1**, 21–26.
- [178] SCOTT, J. & CARRINGTON, P.J. (2011). Introduction. In SCOTT, J., & CARRINGTON, P., (eds.), *The SAGE handbook of social network analysis*, 1–7, SAGE publications. 1, 5, 11, 40
- [179] SEIDMAN, S.B. (1983). Network structure and minimum degree. *Social networks*, **5**, 269–287. 42

- [180] SEIDMAN, S.B. & FOSTER, B.L. (1978). A graph-theoretic generalization of the clique concept*. *Journal of Mathematical sociology*, **6**, 139–154. 85
- [181] SERINO, M. (2013). Theatre provision and decentralization in a region of southern Italy. *New Theatre Quarterly*, **29**, 61–75.
- [182] SERINO, M. (2015). Reti di collaborazione tra teatri di produzione in Campania. *Sociologia del lavoro*, **138**, 121–137. 81, 137
- [183] SERINO, M., D’AMBROSIO, D. & RAGOZINI, G. (2016). *Dynamics of collaboration networks: A multidimensional approach to the study of time-varying two-mode networks in the theatre industry*. paper presented at *DATA SCIENCE & SOCIAL RESEARCH. International Conference*, NAPOLI.
- [184] SIMMEL, G. (1908). *Soziologie: untersuchungen über die formen der vergesellschaftung*. Duncker & Humblot. 45
- [185] SINGH, L., BEARD, M., GETOOR, L. & BLAKE, M.B. (2007). Visual mining of multi-modal social networks at different abstraction levels. In *Information Visualization, 2007. IV’07. 11th International Conference*, 672–679, IEEE.
- [186] SMITH, D.A. & WHITE, D.R. (1992). Structure and dynamics of the global economy: network analysis of international trade 1965–1980. *Social forces*, **70**, 857–893. 96
- [187] SOKAL, R.R. (1958). A statistical method for evaluating systematic relationships. *Univ Kans Sci Bull*, **38**, 1409–1438.
- [188] VICHI, M. & SAPORTA, G. (2009). Clustering and disjoint principal component analysis. *Computational Statistics & Data Analysis*, **53**, 3194–3208.
- [189] WANG, J. & LAI, C. (2008). Detecting groups of similar components in complex networks. *New Journal of Physics*, **10**, 123023.

- [190] WARNER, W.L. & LUNT, P.S. (1941). *The social life of a modern community*. Yale University Press. 2, 8, 11, 19, 20, 21, 26, 28, 41, 42, 44, 71
- [191] WASSERMAN, S. & FAUST, K. (1989). Canonical analysis of the composition and structure of social networks. *Sociological methodology*, **19**, 1–42. 86
- [192] WASSERMAN, S. & FAUST, K. (1994). *Social network analysis: Methods and applications*, vol. 8. Cambridge university press. 5, 6, 8, 10, 11, 17, 18, 27, 30, 31, 33, 39, 40, 42, 43, 44, 45, 46, 47, 56, 85, 96
- [193] WATTS, D.J. & STROGATZ, S.H. (1998). Collective dynamics of "small-world" networks. *nature*, **393**, 440–442. 4
- [194] WELLMAN, H.M. (1999). *Social Structures*. Sage. 4, 5, 6
- [195] WHITE, D.R. & REITZ, K.P. (1983). Graph and semigroup homomorphisms on networks of relations. *Social Networks*, **5**, 193–234. 51
- [196] WHITE, H.C. (1963). *An Anatomy of Kinship*. Prentice Hall. 3
- [197] WHITE, H.C. & BREIGER, R.L. (1975). Pattern across networks. *Society*, **12**, 68–74. 45, 98
- [198] WHITE, H.C., BOORMAN, S.A. & BREIGER, R.L. (1976). Social structure from multiple networks. i. blockmodels of roles and positions. *American journal of sociology*, 730–780. 45, 46, 47, 48, 49
- [199] WILSON, R.J. (1996). *An introduction to graph theory*. Produced through Longman Malaysia, PP. 11, 26, 30
- [200] WINSHIP, C. & MANDEL, M. (1983). Roles and positions: A critique and extension of the blockmodeling approach. *Sociological methodology*, **1984**, 314–344. 46

- [201] ZANI, S. & CERIOLI, A. (2007). *Analisi dei dati e data mining per le decisioni aziendali*. Giuffr  Editore.
- [202] ZHOU, D. & HEINEKEN, E. (2009). The use of metaphors in academic communication: traps or treasures. *Ib rica: Revista de la Asociaci n Europea de Lenguas para Fines Espec ficos (AELFE)*, 23–42. 84
- [203] ZIBERNA, A. (2007). *Generalized blockmodeling of valued network: Doctoral thesis*. Ljubljana University.
- [204] ZIBERNA, A. (2007). Generalized blockmodeling of valued networks. *Social networks*, **29**, 105–126.
- [205] ZIBERNA, A. (2013). Generalized blockmodeling of sparse networks. *Metodoloski zvezki*, **10**, 99–119. 72, 106
- [206]  ZIBERNA, A. (2014). Blockmodeling of multilevel networks. *Social Networks*, **39**, 46–61. 10, 72
- [207] ZIBERNA, A. & LAZEGA, E. (2015). Role sets and division of work at two levels of collective agency: The case of blockmodeling a multilevel (inter-individual and inter-organizational) network. In LAZEGA, E. & SNIJDERS, T. AB., (eds.), *Multilevel Network Analysis for the Social Sciences*, 173–209, SAGE publications.